Oblique projections for direction-of-arrival estimation with prior knowledge

Remy Boyer, Guillaume Bouleux

Oblique Projections for Direction-of-Arrival Estimation With Prior Knowledge

Rémy Boyer and Guillaume Bouleux

***Abstract*—Estimation of directions-of-arrival (DOA) is an important problem in various applications and *a priori* knowledge on the source location is sometimes available. To exploit this information, standard methods are based on the orthogonal projection of the steering manifold onto the noise subspace associated with the *a priori* known DOA. In this paper, we derive and analyze the Cramér-Rao bound associated with this model and in particular we point out the limitations of this approach when the known and unknown DOA are closely spaced and the associated sources are uncorrelated (block-diagonal source covariance). To fill this need, we propose to integrate *a priori* known locations of several sources into the MUSIC algorithm based on oblique projection of the steering manifold. Finally, we show that the proposed approach is able to almost completely cancel the influence of the known DOA on the unknown ones for block-diagonal source covariance and for sufficient signal-to-noise ratio (SNR).**

***Index Terms*—Cramér-Rao bound, MUSIC algorithm, orthogonal and oblique projectors, prior knowledge of DOA.**

I. Introduction

DIRECTIONS-OF-ARRIVAL (DOA) of narrowband sources estimation is one of the central problems in passive radar, sensor sonar, radio-astronomy, and seismology. This problem has received considerable attention in the last 30 years, and a variety of techniques for its solution have been proposed. In practical situations, we have sometimes the knowledge of some *a priori* known subset of the DOA as for instance in a radar application where the emitted signal is backscattered by a number of stationary objects with known positions situated in the radar's viewing field [6], [11]. So, several methods have been proposed to incorporate this prior knowledge into an estimation algorithm. Prior knowledge of DOA can be classified into two families depending if we assume *soft* or *hard* constraints [11]. *Soft* constraints mean that we known approximatively all the DOA. This class of method is known under the name of beamspace methods [20] and has received attention as data reduction methods. The second class of approach incorporates the exact knowledge of a subset of the DOA. This constraint is somewhat more restricting but more interesting gains can be expected. The exact knowledge of DOA allows the deflation of the signal subspace and, thus, to mitigate the influence of the known DOA on the unknown ones.

It is well known that subspace-based methods are very sensitive to over or underestimation of the number of sources and a small error on this parameter leads to a poor accuracy. On the other hand, estimate all the DOA and extract from this set, only the unknown DOA is not easy, especially for closely spaced DOA or/and for low SNRs. So, a specific strategy must be carried out to set up an estimation scheme which extracts only the unknown DOA without bias. In [6] and [11], a constrained-MUSIC algorithm has been presented. The key idea is to orthogonally project the noisy array response onto the noise subspace spanned by the steering vectors associated with the known DOA. Note that, we can find the same principle in some sequential MUSIC algorithms as in [3], [14], and [13]. But, in contrast with our framework, the prior knowledge is uncertain since it is constituted from the previously estimated DOA.

In this paper, we derive and analyze the Cramér-Rao bound (CRB), named Prior-CRB (P-CRB), associated with the orthogonal deflation of the signal subspace. In particular, we show that the prior knowledge of a subset of the DOA leads to a smaller variance for coherent (or highly correlated) sources associated with closely spaced DOA. Another result lies in the fact that if the known and unknown DOA are closely spaced and the associated sources are uncorrelated, the orthogonal deflation cannot help. To fill this need, we advocate that a better scheme for the deflation of the signal subspace is based not only on orthogonal projectors but also on oblique projectors. Based on this principle, we propose to rewrite the MUSIC [16], [19] Least-Squares (LS) criterion in the context of the oblique projector algebra. The resulting LS criterion can be decomposed into the sum of two contributions: the first term is a MUSIC-like criterion and the second one is a corrective function which integrates the prior knowledge.

Note that oblique projectors have received relatively little attention in the literature of signal processing. However, we can find in [9] a first application of oblique projectors to sensor arrays. In [1], Behrens and Scharf propose a very detailed review on this topic and several signal processing oriented applications. More recently, these projection operators have been exploited in [24] and [28] in the context of blind channel identification, in image restoration [27], in noise reduction [8] and in the context of DOA estimation by McCloud and Scharf [12]. Remark that the framework of this latter publication is different to the context of this paper since the authors propose a scaled version of the MUSIC algorithm based on oblique projection but they do not assume prior knowledge of DOA. Finally, our methodology is

Manuscript received September 19, 2006; revised April 27, 2007. This work was presented in part at the Fourth IEEE Workshop on Sensor Array and Multi-Channel Processing (SAM-2006). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Simon J. Godsill.

The authors are wiith the Université Paris XI (UPS), CNRS—LSS-Supélec, 91190 Gif-sur-Yvette, France (e-mail: remy.boyer@lss.supelec.fr; guillaume.bouleux@lss.supelec.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.909348

also valid for all the methods belonging to the class of subspace fitting techniques [26]. In addition, some potential applications of this algorithm can be considered in biomedical signal analysis [5], [22] and harmonic retrieval [25].

After this brief introduction, Section II discusses the matrix-based model of the DOA problem and defines in particular the partitioned steering manifold. In Section III, we present our derivation and the analysis of the P-CRB. In Section IV, we explain how integrate into the MUSIC algorithm the prior knowledge of a subset of the DOA by means of oblique projectors. Section V is dedicated to the practical implementation of the spectral and the root versions of the Prior-MUSIC algorithm. Section VII presents some numerical simulations. Finally, we recall some important facts on oblique projectors and we present the demonstration of the theorems in the Appendix. We denote by bold font vectors and matrices. In addition, T, H, $\dagger$, $\mathrm{Tr}(\cdot), \det(\cdot), \otimes, \bullet$ and $\oplus$ mean, respectively, transpose, conjugated-transpose, Moore-Penrose pseudoinverse, trace, determinant, Kronecker product, Hadamard product, and the direct sum of two subspaces.

II. Matrix-Based Representation of the DOA Estimation Problem

In this section, we introduce the standard matrix-based representation of the DOA estimation problem for an Uniform Linear Array (ULA) and in particular, we define the notion of partitioned steering manifold.

A. Parametric Multi-Input Multi-Output (MIMO) Model

Assume there are M narrowband plane waves simultaneously incident on an ULA with L sensors. The complex array response for the tth snapshot is given by

$$\boldsymbol{y}(t) = \boldsymbol{x}(t) + \sigma \boldsymbol{n}(t) \quad \text{with } \boldsymbol{x}(t) = \boldsymbol{Z}\boldsymbol{\alpha}(t) \tag{1}$$

where $\boldsymbol{y}(t) = [y_1(t) \dots y_L(t)]^T$, $\boldsymbol{\alpha}(t) = [\alpha_1(t) \dots \alpha_M(t)]^T$, $y_\ell(t)$ is the noisy observation on the ℓth sensor and $\alpha_m(t)$ is the complex amplitude of the mth source. The noise vector is denoted by $\boldsymbol{n}(t) = [n_1(t) \dots n_L(t)]^T$ in which the noise on each sensor, denoted by $n_\ell(t)$, is assumed to be additive, white, and Gaussian of parameter $\mathcal{N}(0,1)$ and σ is a positive real parameter. Matrix $\boldsymbol{Z}$ is the $L \times M$ Vandermonde steering manifold defined by

$$\boldsymbol{Z} = [\,\boldsymbol{p}(\theta_1) \quad \boldsymbol{p}(\theta_2) \quad \dots \quad \boldsymbol{p}(\theta_M)\,] \tag{2}$$

where $\boldsymbol{p}(\theta)$ is the steering vector parameterized by DOA θ (in radian), given by

$$\boldsymbol{p}(\theta) = [\,1 \quad e^{-2i\pi(\Delta/c)\sin(\theta)} \quad \cdots \quad e^{-2i\pi(\Delta/c)\sin(\theta)(L-1)}\,]^T$$

in which Δ is the intersensor distance and c is the wavelength. Parameter M is assumed to be known or previously estimated ([18, Appendix C]). So, the parametric MIMO model for T snapshots can be written according to

$$\boldsymbol{Y} = [\,\boldsymbol{y}(1) \quad \dots \quad \boldsymbol{y}(T)\,] = \boldsymbol{X} + \sigma\boldsymbol{N} \tag{3}$$

where $\boldsymbol{X} = [\boldsymbol{x}(1) \dots \boldsymbol{x}(T)] = Z\Lambda$ with $\boldsymbol{\Lambda} = [\boldsymbol{\alpha}(1) \dots \boldsymbol{\alpha}(T)]$ and $\boldsymbol{N} = [\boldsymbol{n}(1) \dots \boldsymbol{n}(T)]$ is the noise matrix.

B. Partitioned Steering Manifold and Deflated Signal Subspace

Assume that we know $M-S$ DOA among M. Without loss of generality, the steering manifold $\boldsymbol{Z}$ can be partitioned according to

$$\boldsymbol{Z} = \left[\left.\underbrace{\boldsymbol{p}(\theta_1) \dots \boldsymbol{p}(\theta_S)}_{\boldsymbol{A}(\text{unknown})}\right| \underbrace{\boldsymbol{p}(\theta_{S+1}) \dots \boldsymbol{p}(\theta_M)}_{\boldsymbol{B}(\text{known})}\right] \tag{4}$$

where the $L \times S$ submanifold A is the matrix composed by the S desired DOA and submanifold B collects the $M-S$ *a priori* known DOA. We name $\mathcal{R}(\boldsymbol{A})$ the subspace of interest or the deflated signal subspace as its dimension is $M-S$ which is smaller than M, the dimension of the signal subspace $\mathcal{R}(\boldsymbol{Z})$. In addition, we assume that the DOA are all distinct ($\theta_i \neq \theta_j$ for $i \neq j$) or equivalently the rank of the steering manifold is M (here, we assume $M \leq L$) then $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}(\boldsymbol{B})$ intersect trivially, i.e., $\mathcal{R}(\boldsymbol{A}) \cap \mathcal{R}(\boldsymbol{B}) = \{\boldsymbol{0}\}$. This implies that $\mathcal{R}(\boldsymbol{Z}) = \mathcal{R}(\boldsymbol{A}) \oplus \mathcal{R}(\boldsymbol{B})$.

C. Structure of the Spatial Covariance

The $(L \times L)$ spatial covariance matrix admits a *Vandermonde-type plus noise* decomposition according to

$$\boldsymbol{R_Y} = \mathbb{E}(\boldsymbol{Y}\boldsymbol{Y}^H) = \boldsymbol{R_X} + \sigma^2 \boldsymbol{I}_L \tag{5}$$

where $\mathbb{E}(\cdot)$ is the mathematical expectation and the noise-free spatial covariance is given by

$$\boldsymbol{R_X} = \boldsymbol{Z}\boldsymbol{R_\Lambda}\boldsymbol{Z}^H. \tag{6}$$

If we assume that all the sources are correlated, the source covariance $\boldsymbol{R_\Lambda}$ is a full matrix, otherwise in the sequel we will sometimes assume that this matrix is block-diagonal, i.e., the sources associated with the known and unknown parts of the steering manifold are uncorrelated. In that case, we have

$$\boldsymbol{R_Y} = \boldsymbol{R_A} + \boldsymbol{R_B} + \sigma^2 \boldsymbol{I}_L \tag{7}$$

with $\boldsymbol{R_A} = \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{\Lambda_A}}\boldsymbol{A}^H$ and $\boldsymbol{R_B} = \boldsymbol{B}\boldsymbol{R}_{\boldsymbol{\Lambda_B}}\boldsymbol{B}^H$ where $\boldsymbol{R}_{\boldsymbol{\Lambda_A}}$ (respectively, $\boldsymbol{R}_{\boldsymbol{\Lambda_B}}$) is the covariance associated with the unknown (respectively, known) sources.

III. Derivation and Analysis of the Prior-CRB

A. Incorporate Prior Knowledge

In [6] and [11], a prior knowledge MUSIC algorithm has been introduced and analyzed. This algorithm, called constrained MUSIC, is based on the projection of the noisy array response onto $\mathcal{R}(\boldsymbol{B})^\perp$. In a view to derive the Prior-CRB, we "vectorize" model (3) according to

$$\boldsymbol{y} = \mathrm{vec}(\boldsymbol{Y}) = [\,\boldsymbol{y}(1)^T \quad \dots \quad \boldsymbol{y}(T)^T\,]^T = \boldsymbol{x} + \sigma\boldsymbol{n} \tag{8}$$

in which the noise is denoted by vector $\boldsymbol{n} = \mathrm{vec}(\boldsymbol{N})$. Consequently, $\sigma\boldsymbol{n}$ is an additive white Gaussian process of parameters $\mathcal{N}(0, \sigma^2 \boldsymbol{I}_{LT})$ and

$$\begin{aligned}\boldsymbol{x} &= \mathrm{vec}(\boldsymbol{X}) \\ &= [\,\boldsymbol{x}(1)^T \quad \dots \quad \boldsymbol{x}(T)^T\,]^T = (\boldsymbol{I}_T \otimes \boldsymbol{Z})\boldsymbol{\lambda} \end{aligned} \tag{9}$$

and $\boldsymbol{\lambda} = [\boldsymbol{\alpha}(1)^T \dots \boldsymbol{\alpha}(T)^T]^T$.

To incorporate prior knowledge of the known DOA set: $\{\theta_{S+1}, \dots, \theta_M\}$ collected in submanifold $\boldsymbol{B}$, model (8) is modified according to

$$\begin{aligned}\boldsymbol{y}_{(\text{prior})} &= \left(\boldsymbol{I}_T \otimes \boldsymbol{P}_{\boldsymbol{B}}^{\perp}\right)\boldsymbol{y} \\ &= \boldsymbol{x}_{(\text{prior})} + \sigma\left(\boldsymbol{I}_T \otimes \boldsymbol{P}_{\boldsymbol{B}}^{\perp}\right)\boldsymbol{n}. \end{aligned} \tag{10}$$

where $\boldsymbol{P}_{\boldsymbol{B}}^{\perp} = \boldsymbol{I}_L - \boldsymbol{B}\boldsymbol{B}^{\dagger}$ and the notation "prior" indicates that we have projected signal $\boldsymbol{y}$ onto the noise subspace associated with the *a priori* known DOA. Following the same notation, we define the noise-free signal according to

$$\boldsymbol{x}_{(\text{prior})} = \left(\boldsymbol{I}_T \otimes \left(\boldsymbol{P}_{\boldsymbol{B}}^{\perp}\boldsymbol{A}\right)\right)\bar{\boldsymbol{\lambda}} \tag{11}$$

where $\bar{\boldsymbol{\lambda}} = [\bar{\boldsymbol{\alpha}}(1)^T \dots \bar{\boldsymbol{\alpha}}(T)^T]^T$ with $\bar{\boldsymbol{\alpha}}(t) = [\alpha_1(t) \dots \alpha_S(t)]^T$.

B. Prior-CRB Based on Model (10) (P-CRB)

There are several ways to derive the P-CRB. A first approach could be to derive this bound in the framework of Gaussian colored processes where the inverse of the noise error covariance, involved in the general CRB formula [21], is computed from a truncated pseudoinverse of projector $\boldsymbol{P}_{\boldsymbol{B}}^{\perp}$. An equivalent and more handy way is based on the following proposition.

Proposition 1: The "compressed"[1] signal

$$\tilde{\boldsymbol{y}}_{(\text{prior})} = \left(\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}^H\right)\boldsymbol{y}_{(\text{prior})} \tag{12}$$

where $\boldsymbol{U}_{\boldsymbol{B}}$ is a $L \times (L - M + S)$ unitary basis of $\mathcal{R}(\boldsymbol{B})^{\perp}$, follows a Gaussian distribution of parameters $\mathcal{N}(\tilde{\boldsymbol{x}}_{(\text{prior})}, \tilde{\boldsymbol{\Gamma}} = \sigma^2 \boldsymbol{I}_{LT})$ where $\tilde{\boldsymbol{x}}_{(\text{prior})} = (\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}^H)\boldsymbol{x}_{(\text{prior})}$.

Proof: See Appendix B.

Let define the signal plus nuisance model parameter vector by $\boldsymbol{\chi} = [\boldsymbol{\chi}'^T \ \sigma^2]^T$ where $\boldsymbol{\chi}' = [\boldsymbol{\theta}^{(S)^T} \ \bar{\boldsymbol{\lambda}}_{\mathcal{R}}^T \ \bar{\boldsymbol{\lambda}}_{\mathcal{I}}^T]^T$, $\boldsymbol{\theta}^{(S)} = [\theta_1 \dots \theta_S]^T$, $\bar{\boldsymbol{\lambda}}_{\mathcal{R}} = \Re e\{\bar{\boldsymbol{\lambda}}\}$ and $\bar{\boldsymbol{\lambda}}_{\mathcal{I}} = \Im m\{\bar{\boldsymbol{\lambda}}\}$. A standard result ([18], Appendix B) is that the Mean Squares Error (MSE) for any unbiased estimate, $\hat{\boldsymbol{\chi}}$, of the parameter vector $\boldsymbol{\chi}$ satisfies $\mathrm{MSE}(\hat{\boldsymbol{\chi}}) = \mathbb{E}((\boldsymbol{\chi} - \hat{\boldsymbol{\chi}})(\boldsymbol{\chi} - \hat{\boldsymbol{\chi}})^H) \geq \mathrm{P-CRB}(\boldsymbol{\chi})$ with

$$\begin{aligned}&\mathrm{P{-}CRB}^{-1}(\boldsymbol{\chi}) \\ &\quad = \mathbb{E}\left(\frac{\partial \ln p(\tilde{\boldsymbol{y}}_{(\text{prior})}, \boldsymbol{\chi})}{\partial \boldsymbol{\chi}} \frac{\partial \ln p(\tilde{\boldsymbol{y}}_{(\text{prior})}, \boldsymbol{\chi})}{\partial \boldsymbol{\chi}}^H\right) \end{aligned} \tag{13}$$

where $p(\boldsymbol{y}, \boldsymbol{\chi})$ is the likelihood function. Consequently, the P-CRB is a lower bound on the minimal achievable variance. According to proposition 1, (13) can be rewritten for Gaussian distribution according to

$$\begin{aligned}&[\mathrm{P-CRB}^{-1}(\boldsymbol{\chi})]_{ij} \\ &\quad = \mathrm{Tr}\left(\tilde{\boldsymbol{\Gamma}}^{-1}\frac{\partial \tilde{\boldsymbol{\Gamma}}}{\partial \boldsymbol{\chi}_i}\tilde{\boldsymbol{\Gamma}}^{-1}\frac{\partial \tilde{\boldsymbol{\Gamma}}}{\partial \boldsymbol{\chi}_j}\right) \\ &\qquad + 2\Re e\left\{\left(\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \boldsymbol{\chi}_i}\right)^H \tilde{\boldsymbol{\Gamma}}^{-1}\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \boldsymbol{\chi}_j}\right\} \end{aligned} \tag{14}$$

where $\boldsymbol{\chi}_i$ denotes the ith block of vector $\boldsymbol{\chi}$ for $i \in [1:4]$. Finally, we can formulate the following result.

Theorem 1: The P-CRB with respect to the "signal" parameter is given by

$$\mathrm{P-CRB}(\boldsymbol{\theta}^{(S)}) = \frac{\sigma^2}{2T}\left[\Re e\left\{\left[\boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}\boldsymbol{D}_{\boldsymbol{A}}\right] \bullet \boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}}\right\}\right]^{-1} \tag{15}$$

where $\boldsymbol{D}_{\boldsymbol{A}} = [\boldsymbol{p}'(\theta_1) \dots \boldsymbol{p}'(\theta_S)]$ with $\boldsymbol{p}'(\theta_k) = \frac{\partial \boldsymbol{p}(\theta)}{\partial \theta}|_{\theta = \theta_k}$ and $\boldsymbol{P}_{\boldsymbol{Z}}^{\perp} = \boldsymbol{I} - \boldsymbol{Z}\boldsymbol{Z}^{\dagger}$.

Proof: See Appendix C.

C. CRB Without Prior Knowledge Based on Model (8)

The derivation of these bounds is well known [18, p. 392, (B.6.32)]. Then, we have

$$\mathrm{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)}) = \frac{\sigma^2}{2T}\left[\Re e\left\{\left[\boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{P}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}_{\boldsymbol{A}}\right] \bullet \boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}}\right\}\right]^{-1} \tag{16}$$

$$\mathrm{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(M)}) = \frac{\sigma^2}{2T}\left[\Re e\left\{\left[\boldsymbol{D}_{\boldsymbol{Z}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}\boldsymbol{D}_{\boldsymbol{Z}}\right] \bullet \boldsymbol{R}_{\boldsymbol{\Lambda}}\right\}\right]^{-1} \tag{17}$$

where $\boldsymbol{\theta}^{(M)} = [\theta_1 \dots \theta_M]^T$, $\boldsymbol{P}_{\boldsymbol{A}}^{\perp} = \boldsymbol{I}_L - \boldsymbol{A}\boldsymbol{A}^{\dagger}$, and $\boldsymbol{D}_{\boldsymbol{Z}} = [\boldsymbol{p}'(\theta_1) \dots \boldsymbol{p}'(\theta_M)]$. Expressions (16) and (17) are, respectively, the CRB over the subspace of interest $\mathcal{R}(\boldsymbol{A})$ and over the whole space $\mathcal{R}(\boldsymbol{Z})$. Remark that the P-CRB mixes the $\mathrm{CRB}_{\boldsymbol{A}}$ and the $\mathrm{CRB}_{\boldsymbol{Z}}$ in the sense that only the unknown derivative steering vectors are projected onto the orthogonal complement of the whole space.

D. Comparison of the Bounds

The following theorem discusses some properties linking the derived bounds.

Theorem 2: For any unbiased estimator of $\boldsymbol{\theta}^{(S)}$, we have the following relations.

- For full $\boldsymbol{R}_{\boldsymbol{\Lambda}}$:
  - (i) $\mathrm{P{-}CRB}(\boldsymbol{\theta}^{(S)}) \geq \mathrm{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$
  - (ii) $\mathrm{P{-}CRB}(\theta_m), \mathrm{CRB}_{\boldsymbol{A}}(\theta_m), \mathrm{CRB}_{\boldsymbol{Z}}(\theta_m) \xrightarrow{L \gg 1} \frac{6}{L^3 T}\frac{\sigma^2}{[\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}}]_{mm}}$, for $m \in [1:S]$.
- If $\boldsymbol{R}_{\boldsymbol{\Lambda}}$ is block-diagonal:
  - (iii) $\mathrm{P{-}CRB}(\boldsymbol{\theta}^{(S)}) = \mathrm{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(S)})$
  - (iv) $\mathrm{P-CRB}(\boldsymbol{\theta}^{(S)}) \gg \mathrm{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$ for closely spaced DOA
  - (v) $\mathrm{P{-}CRB}(\boldsymbol{\theta}^{(S)}) < \mathrm{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(S)})$ for closely spaced DOA and $\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}}$ deficient.
  - (vi) $\mathrm{P{-}CRB}(\boldsymbol{\theta}^{(S)}) \approx \mathrm{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$ for widely spaced DOA.

Proof: See Appendix D.

[1] In the sense that $(\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}^H)$ is a "fat" matrix (more columns than rows).

Note that *closely spaced DOA* means that at least one known DOA is close to at least one unknown DOA.[2]

If the sources are correlated and according to property (i), we cannot expect perform better than the CRB over the subspace of interest, $\mathcal{R}(\boldsymbol{A})$. This is a fundamental limit. Regarding property (ii), prior knowledge cannot help for a large number of sensors. If the known and unknown sources are uncorrelated, the known sources are not coherent and the DOA are widely spaced, then properties (iii) and (vi) mean that all the CRB are merged. Inversely [*cf.* property (v)], if the known sources are coherent or highly correlated, prior knowledge can be exploited. In conclusion, the use of orthogonal projector to introduce prior knowledge into an estimation algorithm is recommended for some limit situations as for closely spaced DOA associated with coherent (or possibly highly correlated) sources with small/moderate number of sensors. Some of these conclusions have been already obtained in [6] and [11] in the context of the constrained MUSIC algorithm but here we present a more general argumentation since we base our analysis on a statistical quantity which is independent of the specific choice of the estimation algorithm. An important point is that, for finite L, property (iv) suggests that the orthogonal projector does not completely cancel the influence of the known DOA on the estimation of the unknown ones and does not reach the CRB over $\mathcal{R}(\boldsymbol{A})$ for block-diagonal source covariance associated with closely spaced DOA. In Section IV, we propose an estimation scheme which solves this problem.

IV. Prior Knowledge-Based MUSIC Algorithms

To be in line with property (iv), we assume that the sources associated with the known and with the unknown parts of the steering manifold are uncorrelated. Consequently, the spatial covariance matrix of the sources is block-diagonal and is defined in (7). In addition, we denote unknown quantities by the "hat" symbol. According to this notation, we have $\hat{\boldsymbol{Z}} = [\hat{\boldsymbol{A}}\ \boldsymbol{B}]$.

It can be seen that the Constrained-MUSIC (CMUSIC) algorithm [6], [11] attempts to find one component at a time which is most orthogonal to the noise subspace of the partially known steering manifold, $\hat{\boldsymbol{Z}}$. The CMUSIC optimization problem can be described according to $\arg\min_\theta f_{\mathrm{CMUSIC}}(\theta)$ where

$$f_{\mathrm{CMUSIC}}(\theta) = \left\|\boldsymbol{P}^{\perp}_{\hat{\boldsymbol{Z}}}\boldsymbol{p}(\theta)\right\|^2 = \left\|\boldsymbol{P}^{\perp}_{[\hat{\boldsymbol{A}}\ \boldsymbol{B}]}\boldsymbol{p}(\theta)\right\|^2. \tag{18}$$

Note that this problem is different to the standard unconstrained problem associated with the MUSIC algorithm: $\arg\min_\theta f_{\mathrm{MUSIC}}(\theta)$ with $f_{\mathrm{MUSIC}}(\theta) = \|\boldsymbol{P}^{\perp}_{[\hat{\boldsymbol{A}}\ \hat{\boldsymbol{B}}]}\boldsymbol{p}(\theta)\|^2$ where both matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are unknown.

Following the formalism introduced in [7], [16] for the MUSIC algorithm, the constrained Prior-MUSIC criterion is given by

$$\arg\min_{\boldsymbol{\alpha},\theta} \mathcal{C}(\boldsymbol{\alpha},\theta) \quad \text{subject to } \boldsymbol{p}(\theta) \in \mathcal{R}(\hat{\boldsymbol{A}}) \tag{19}$$

where the cost function is defined by

$$\mathcal{C}(\boldsymbol{\alpha},\theta) = \left\|\boldsymbol{p}(\theta) - \boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\hat{\boldsymbol{Z}}\boldsymbol{\alpha}\right\|^2 \tag{20}$$

in which $\boldsymbol{\alpha}$ is a complex amplitude vector. Let $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_{\hat{\boldsymbol{A}}}^T\ \boldsymbol{\alpha}_{\boldsymbol{B}}^T]^T$ then the cost function $\mathcal{C}(\boldsymbol{\alpha},\theta)$ can be rewritten according to

$$\begin{aligned}\mathcal{C}(\boldsymbol{\alpha},\theta) &= \left\|\boldsymbol{p}(\theta) - \boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}(\hat{\boldsymbol{A}}\boldsymbol{\alpha}_{\hat{\boldsymbol{A}}} + \boldsymbol{B}\boldsymbol{\alpha}_{\boldsymbol{B}})\right\|^2 &(21)\\ &= \|\boldsymbol{p}(\theta) - \hat{\boldsymbol{A}}\boldsymbol{\alpha}_{\hat{\boldsymbol{A}}}\|^2 &(22)\end{aligned}$$

since $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\hat{\boldsymbol{A}} = \hat{\boldsymbol{A}}$ and $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\boldsymbol{B} = \boldsymbol{0}$. Consequently according to (22), minimizing $\mathcal{C}(\boldsymbol{\alpha},\theta)$ is equivalent to look for vectors $\boldsymbol{p}(\theta)$ in the subspace of interest $\mathcal{R}(\hat{\boldsymbol{A}})$ and in the same time to discard the DOA belonging to the known space $\mathcal{R}(\boldsymbol{B})$. In contrast with orthogonal projectors, the steering manifold $\hat{\boldsymbol{A}}$ is not affected by projector $\boldsymbol{P}^{\perp}_{\boldsymbol{B}}$ [cf. (11)].

A. Prior-MUSIC (P-MUSIC) Algorithm

A standard minimal-norm solution of $\mathcal{C}(\boldsymbol{\alpha},\theta)$ with respect to parameter $\boldsymbol{\alpha}$ is given by

$$\begin{aligned}&\tilde{\boldsymbol{\alpha}} = \begin{bmatrix}\tilde{\boldsymbol{\alpha}}_{\hat{\boldsymbol{A}}}\\ \tilde{\boldsymbol{\alpha}}_{\boldsymbol{B}}\end{bmatrix} = \left(\boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\hat{\boldsymbol{Z}}\right)^{\dagger}\boldsymbol{p}(\theta) = \begin{bmatrix}\hat{\boldsymbol{A}}^{\dagger}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{p}(\theta)\\ &\Longrightarrow \begin{cases}\tilde{\boldsymbol{\alpha}}_{\hat{\boldsymbol{A}}} = \hat{\boldsymbol{A}}^{\dagger}\boldsymbol{p}(\theta)\\ \tilde{\boldsymbol{\alpha}}_{\boldsymbol{B}} = \boldsymbol{0}.\end{cases}\end{aligned} \tag{23}$$

This solution satisfies $\forall \boldsymbol{p}(\theta) \in \mathcal{R}(\boldsymbol{Z})$, $\frac{\partial \mathcal{C}(\boldsymbol{\alpha},\theta)}{\partial \boldsymbol{\alpha}}|_{\boldsymbol{\alpha}=\tilde{\boldsymbol{\alpha}}} = \boldsymbol{0}$. Consequently, the cost function of the P-MUSIC algorithm

$$\mathcal{C}(\tilde{\boldsymbol{\alpha}},\theta) = \left\|\boldsymbol{P}^{\perp}_{\hat{\boldsymbol{A}}}\boldsymbol{p}(\theta)\right\|^2 \tag{24}$$

reaches its minimum value (*wrt.* $\boldsymbol{\alpha}$) for the known and unknown DOA. However, criterion (19) must be minimal only for the known DOA. So, we can expect that in some limit situations as for low SNR or for closely spaced DOA, this approach will be suboptimal. Consequently, in the next paragraph, we propose another approach which completely solves criterion (19).

B. Weighted Prior-MUSIC (WP-MUSIC) Algorithm

1) A Second Resolution Based on the Obliquely Weighted Pseudoinverse: To solve criterion (19), consider the following minimal-norm solution:

$$\bar{\boldsymbol{\alpha}} = \hat{\boldsymbol{Z}}^{\ddagger}\boldsymbol{p}(\theta) \tag{25}$$

where $\hat{\boldsymbol{Z}}^{\ddagger}$ is the *obliquely weighted* pseudoinverse[3] defined by $\hat{\boldsymbol{Z}}^{\ddagger} = \hat{\boldsymbol{Z}}^{\dagger}\boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}$. Then, the cost function can be rewritten according to

$$\begin{aligned}\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta) &= \left\|\boldsymbol{p}(\theta) - \boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\boldsymbol{P}_{\hat{\boldsymbol{Z}}}\boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\boldsymbol{p}(\theta)\right\|^2 &(26)\\ &= \left\|\left(\boldsymbol{I}_L - \boldsymbol{E}_{(\hat{\boldsymbol{A}}\ \boldsymbol{B})}\right)\boldsymbol{p}(\theta)\right\|^2 &(27)\\ &= \left\|\left(\boldsymbol{P}^{\perp}_{\hat{\boldsymbol{Z}}} + \boldsymbol{E}_{(\boldsymbol{B}\ \hat{\boldsymbol{A}})}\right)\boldsymbol{p}(\theta)\right\|^2. &(28)\end{aligned}$$

[2]More precisely, the inter-DOA distance must be under the Rayleigh resolution [23], i.e., $|\pi(\sin(\theta_2) - \sin(\theta_1))| \ll 2\sqrt{6/(L^2-1)}$.

[3]We do not confuse the *obliquely weighted* pseudoinverse with the standard *oblique* pseudoinverse defined in [1].

The previous expressions are derived by using some basic properties of oblique projectors defined in (53) and (54) and the fact that projectors are idempotent. So, the final WP-MUSIC criterion is given by $\arg\min_\theta \mathcal{C}(\bar{\boldsymbol{\alpha}},\theta)$. Note that

- $\forall \boldsymbol{p}(\theta) \in \mathcal{R}(\hat{\boldsymbol{A}})$, we have $\boldsymbol{P}_{\hat{\boldsymbol{Z}}}^{\perp}\boldsymbol{p}(\theta) = \boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\boldsymbol{p}(\theta) = 0$. So, $\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta) = 0$ and thus is minimal.
- $\forall \boldsymbol{p}(\theta) \in \mathcal{R}(\boldsymbol{B})$, we have $\boldsymbol{P}_{\hat{\boldsymbol{Z}}}^{\perp}\boldsymbol{p}(\theta) = 0$ and $\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\boldsymbol{p}(\theta) = \boldsymbol{p}(\theta)$ and thus $\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta) = L$. So, $\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta)$ is not minimal.

To show that $\bar{\boldsymbol{\alpha}}$ is a minimal-norm solution of criterion (19), we consider the partial derivative *wrt.* $\boldsymbol{\alpha} = \bar{\boldsymbol{\alpha}}$ of the cost function according to

$$\left.\frac{\partial \mathcal{C}(\boldsymbol{\alpha},\theta)}{\partial \boldsymbol{\alpha}^H}\right|_{\boldsymbol{\alpha}=\bar{\boldsymbol{\alpha}}} = -\boldsymbol{Z}^H \boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}^H \left(\boldsymbol{P}_{\hat{\boldsymbol{Z}}}^{\perp} + \boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\right)\boldsymbol{p}(\theta). \tag{29}$$

We have two cases,

- $\forall \boldsymbol{p}(\theta) \in \mathcal{R}(\hat{\boldsymbol{A}})$ then $\frac{\partial \mathcal{C}(\boldsymbol{\alpha},\theta)}{\partial \boldsymbol{\alpha}}|_{\boldsymbol{\alpha}=\bar{\boldsymbol{\alpha}}} = \boldsymbol{0}$. So, $\bar{\boldsymbol{\alpha}}$ is a minimal-norm solution of criterion (19) for the unknown DOA.
- $\forall \boldsymbol{p}(\theta) \in \mathcal{R}(\boldsymbol{B})$ then $\frac{\partial \mathcal{C}(\boldsymbol{\alpha},\theta)}{\partial \boldsymbol{\alpha}}|_{\boldsymbol{\alpha}=\bar{\boldsymbol{\alpha}}} \neq \boldsymbol{0}$. Thus for the known DOA, $\bar{\boldsymbol{\alpha}}$ is not a minimal-norm solution or in other words, the cost function $\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta)$ does not reached its minimum *wrt.* $\boldsymbol{\alpha}$ for the known DOA. This fact is very desirable since criterion (19) must be minimum *only* for the unknown DOA. This property can be understood as a reinforcement of the rejection of the known DOA. In addition, for the known DOA, it comes $\bar{\boldsymbol{\alpha}} = \boldsymbol{0}$.

2) Link To the MUSIC-Like Criterion: One can easily verify that $\boldsymbol{P}_{\hat{\boldsymbol{Z}}}^{\perp}\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})} = \boldsymbol{0}$ and, therefore, the cost function $\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta)$ can be decomposed into two contributions according to

$$\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta) = f_{\text{CMUSIC}}(\theta) + f_{\text{COR}}(\theta) \tag{30}$$

where $f_{\text{CMUSIC}}(\theta)$ has been defined in (18) and $f_{\text{COR}}(\theta) = \|\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\boldsymbol{p}(\theta)\|^2$. As we can see, the above expression is a CMUSIC criterion with an additional corrective term which takes into account the prior knowledge.

C. Large Number of Sensors

We begin by exposing an asymptotic result regarding oblique projectors.

1) Proposition 2: For large L and if $\mathcal{R}(\hat{\boldsymbol{A}})$ and $\mathcal{R}(\boldsymbol{B})$ intersect trivially, we have $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})} \xrightarrow{L\gg 1} \boldsymbol{P}_{\hat{\boldsymbol{A}}} = (1/L)\hat{\boldsymbol{A}}\hat{\boldsymbol{A}}^H$ and $\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})} \xrightarrow{L\gg 1} \boldsymbol{P}_{\boldsymbol{B}} = (1/L)\boldsymbol{B}\boldsymbol{B}^H$.

Proof: As $\mathcal{R}(\hat{\boldsymbol{A}}) \cap \mathcal{R}(\boldsymbol{B}) = \{\boldsymbol{0}\}$, we have $\boldsymbol{p}(\theta_i)^H\boldsymbol{p}(\theta_j) \xrightarrow{L\gg 1} 0$ where $i \in [1:S]$ and $j \in [S+1:M]$ then $\mathcal{R}(\hat{\boldsymbol{A}})$ and $\mathcal{R}(\boldsymbol{B})$ are mutually orthogonal. This implies $(1/L)\boldsymbol{A}^H\boldsymbol{B},\ (1/L)\boldsymbol{B}^H\boldsymbol{A} \xrightarrow{L\gg 1} \boldsymbol{0}$. Using these properties together with the definition of an oblique projector given in (52), it is easy to show the proposition.

It is straightforward to see that for a large number of sensors, we have

$$\bar{\boldsymbol{\alpha}},\ \tilde{\boldsymbol{\alpha}} \xrightarrow{L\gg 1} \boldsymbol{\alpha}^\star \quad \text{where } \boldsymbol{\alpha}^\star = \frac{1}{L}\begin{bmatrix}\hat{\boldsymbol{A}}^H \\ \boldsymbol{0}\end{bmatrix}\boldsymbol{p}(\theta). \tag{31}$$

According to (23) and the fact that $\boldsymbol{A}^\dagger \xrightarrow{L\gg 1} \frac{1}{L}\boldsymbol{A}^H$, it is direct to show $\tilde{\boldsymbol{\alpha}} \xrightarrow{L\gg 1} \boldsymbol{\alpha}^\star$. In addition, it comes

$$\begin{aligned}
\bar{\boldsymbol{\alpha}} &= \boldsymbol{Z}^\dagger \boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}\boldsymbol{p}(\theta) \\
&= (\boldsymbol{Z}^H\boldsymbol{Z})^{-1}\boldsymbol{Z}^H\hat{\boldsymbol{A}}(\hat{\boldsymbol{A}}^H\boldsymbol{P}_{\boldsymbol{B}}^{\perp}\hat{\boldsymbol{A}})^{-1}\hat{\boldsymbol{A}}^H\boldsymbol{P}_{\boldsymbol{B}}^{\perp}\boldsymbol{p}(\theta) && (32)\\
&\xrightarrow{L\gg 1} \frac{L}{L^2}\left(\boldsymbol{I}_M\begin{bmatrix}\boldsymbol{I}_S \\ \boldsymbol{0}\end{bmatrix}\boldsymbol{I}_S\right)\boldsymbol{A}^H\boldsymbol{p}(\theta) && (33)\\
&= \frac{1}{L}\begin{bmatrix}\hat{\boldsymbol{A}}^H \\ \boldsymbol{0}\end{bmatrix}\boldsymbol{p}(\theta) && (34)\\
&= \boldsymbol{\alpha}^\star && (35)
\end{aligned}$$

where we have used the results given in proposition 2. Consequently, the P-MUSIC and WP-MUSIC cost functions become

$$\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta), \mathcal{C}(\tilde{\boldsymbol{\alpha}},\theta) \xrightarrow{L\gg 1} \mathcal{C}(\boldsymbol{\alpha}^\star,\theta) = \left\|\boldsymbol{P}_{\hat{\boldsymbol{A}}}^{\perp}\boldsymbol{p}(\theta)\right\|^2.$$

This result means that asymptotically, the criterions of the P-MUSIC and WP-MUSIC are in fact the criterion of the MUSIC over the *orthonormalized* subspace of interest. In addition, the WP-MUSIC and P-MUSIC algorithms based, respectively, on criterion $\mathcal{C}(\bar{\boldsymbol{\alpha}},\theta)$ and $\mathcal{C}(\tilde{\boldsymbol{\alpha}},\theta)$ are asymptotically equivalent. However, for more realistic situations where L takes moderate values, we show in the simulation part that the two approaches are no longer equivalent, as expected in Sections IV-A and IV-B.

V. Implementation of the WP-MUSIC Criterion

As claimed at the end of Section IV-B, we focus our analysis on the WP-MUSIC algorithm. To implement this algorithm, we have two possibilities. We can use (27) or (28). The latter is more readable since it explains how the WP-MUSIC algorithm works. However, it is preferable to implement (27) for the two following reasons.

1) More DOA can be estimated. Since the second expression of the WP-MUSIC criterion involves projector $\boldsymbol{P}_{\hat{\boldsymbol{Z}}}^{\perp}$, we have to satisfy constraint $M \leq L$ to ensure that $\hat{\boldsymbol{Z}}$ is a rank-M matrix. For the first expression, only matrices $\hat{\boldsymbol{A}}$ and $\boldsymbol{B}$ are involved through projector $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$. In that case, we have to satisfy the following constraints:

$$\begin{gathered}
\mathcal{R}(\hat{\boldsymbol{A}}) \subseteq \mathcal{R}(\hat{\boldsymbol{Z}}) \Longrightarrow S \leq M,\\
\hat{\boldsymbol{A}} \text{ is of rank-}S \Longleftrightarrow S \leq L,\\
\boldsymbol{B} \text{ is of rank-}(M-S) \Longleftrightarrow M-S \leq L.
\end{gathered}$$

The two last constraints can be reformulated as $\max(S, M-S) \leq L$ which is less restrictive than $M \leq L$ since we have $\max(S, M-S) \leq M$. In fact, combining the constraints on the rank of matrices $\hat{A}$ and B, we obtain $M \leq 2L$ which allows possible values for M greater than L.

2) In a computational point of view, (28) involves the estimation of projectors $\boldsymbol{P}_{\hat{\boldsymbol{Z}}}^{\perp}$ and $\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}$ while (27) involves only the estimation of projector $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$.

A. Estimation of Oblique Projectors

1) Invariant to Change of Basis: The oblique projectors $\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}$ and $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$ are invariant to change of basis. Indeed a basis of space $\mathcal{R}(\hat{\boldsymbol{A}})$ is not unique, so consider another basis $\boldsymbol{\Phi}$ such as $\mathcal{R}(\boldsymbol{\Phi}) = \mathcal{R}(\hat{\boldsymbol{A}})$. We know that there exits an invertible matrix $\boldsymbol{\Theta}$ such as $\boldsymbol{\Phi}\boldsymbol{\Theta} = \hat{\boldsymbol{A}}$. In that case, it comes the two following equalities $\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})} = \boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}}\boldsymbol{\Theta}^{-1})}$ and $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})} = \boldsymbol{E}_{(\hat{\boldsymbol{A}}\boldsymbol{\Theta}^{-1}\,\boldsymbol{B})}$. This invariance property for $\boldsymbol{E}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}$ is a consequence of the fact that $P^{\perp}_{\hat{\boldsymbol{A}}}$ is essentially unique since $P_{\hat{\boldsymbol{A}}} = \hat{\boldsymbol{A}}\boldsymbol{\Theta}^{-1}\boldsymbol{\Theta}\hat{\boldsymbol{A}}^{\dagger} = \hat{\boldsymbol{A}}\hat{\boldsymbol{A}}^{\dagger}$. For projector $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$, we can show this result in the following manner:

$$\boldsymbol{E}_{(\hat{\boldsymbol{A}}\boldsymbol{\Theta}^{-1}\,\boldsymbol{B})} = \hat{\boldsymbol{A}}\boldsymbol{\Theta}^{-1}\boldsymbol{\Theta}\left(\hat{\boldsymbol{A}}^{H}\boldsymbol{P}^{\perp}_{\boldsymbol{B}}\hat{\boldsymbol{A}}\right)^{-1} \times(\boldsymbol{\Theta}^{-1}\boldsymbol{\Theta})^{H}\hat{\boldsymbol{A}}^{H}\boldsymbol{P}^{\perp}_{\boldsymbol{B}} = \boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}. \quad (36)$$

2) Estimation of Projector $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$: In criterion (27), we need to partially estimate projector $\boldsymbol{E}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$. So, knowing projector $\boldsymbol{P}^{\perp}_{\boldsymbol{B}}$, we have to estimate a basis of $\mathcal{R}(\hat{\boldsymbol{A}})$. Consider the sample weighted spatial covariance matrix of the noise-free array response:

$$\bar{\boldsymbol{R}}_{\boldsymbol{X}} = \frac{1}{T}\boldsymbol{P}^{\perp}_{\boldsymbol{B}}\boldsymbol{X}\boldsymbol{X}^{H}. \quad (37)$$

As $\boldsymbol{R}_{\boldsymbol{X}}$ admits a Vandermonde-type decomposition according to (6) and as we assume that $\boldsymbol{R}_{\boldsymbol{\Lambda}}$ is block-diagonal, it comes

$$\bar{\boldsymbol{R}}_{\boldsymbol{X}} = \boldsymbol{P}^{\perp}_{\boldsymbol{B}}\hat{\boldsymbol{Z}}\boldsymbol{R}_{\boldsymbol{\Lambda}}\hat{\boldsymbol{Z}}^{H} = \boldsymbol{P}^{\perp}_{\boldsymbol{B}}\hat{\boldsymbol{A}}\boldsymbol{R}_{\boldsymbol{\Lambda}_{\hat{A}}}\hat{\boldsymbol{A}}^{H} \quad (38)$$

and the rank of $\bar{\boldsymbol{R}}_{\boldsymbol{X}}$ is S. Now, consider the Singular Value Decomposition (SVD, [18, p. 355]) of the sample weighted spatial covariance

$$\bar{\boldsymbol{R}}_{\boldsymbol{X}} = \boldsymbol{U}\boldsymbol{\Sigma}\left[\boldsymbol{V}_{\hat{\boldsymbol{A}}} \quad \bar{\boldsymbol{V}}_{\hat{\boldsymbol{A}}}\right]^{H} \quad (39)$$

where $\boldsymbol{V}_{\hat{\boldsymbol{A}}}$ is a $L\times S$ unitary basis of $\mathcal{R}(\hat{\boldsymbol{A}})$ and $\bar{\boldsymbol{V}}_{\hat{\boldsymbol{A}}}$ is a $L\times(L-S)$ unitary basis of the null-space of the sample weighted spatial covariance, denoted by $\mathcal{R}(\hat{\boldsymbol{A}})^{\perp}$. We use the right basis since projector $\boldsymbol{P}^{\perp}_{\boldsymbol{B}}$ destroys the Hermitian character of the sample spatial covariance. Projector $\hat{\boldsymbol{E}}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$ can be computed according to

$$\begin{aligned}\hat{\boldsymbol{E}}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})} &= \boldsymbol{V}_{\hat{\boldsymbol{A}}}\left(\boldsymbol{V}^{H}_{\hat{\boldsymbol{A}}}\boldsymbol{P}^{\perp}_{\boldsymbol{B}}\boldsymbol{V}_{\hat{\boldsymbol{A}}}\right)^{-1}\boldsymbol{V}^{H}_{\hat{\boldsymbol{A}}}\boldsymbol{P}^{\perp}_{\boldsymbol{B}} && (40)\\ &= \boldsymbol{V}_{\hat{\boldsymbol{A}}}(\boldsymbol{P}^{\perp}_{\boldsymbol{B}}\boldsymbol{V}_{\hat{\boldsymbol{A}}})^{\dagger} && (41)\\ &= \left(\boldsymbol{P}^{\perp}_{\boldsymbol{B}}\hat{\boldsymbol{P}}_{\hat{\boldsymbol{A}}}\right)^{\dagger} && (42)\end{aligned}$$

where $\hat{\boldsymbol{P}}_{\hat{\boldsymbol{A}}} = \boldsymbol{V}_{\hat{\boldsymbol{A}}}\boldsymbol{V}^{H}_{\hat{\boldsymbol{A}}}$. In presence of noise, it is preferable to consider $\bar{\boldsymbol{R}}_{Y} = (1/T)\boldsymbol{U}^{H}_{\boldsymbol{B}}\boldsymbol{Y}\boldsymbol{Y}^{H}$ where $\boldsymbol{U}_{\boldsymbol{B}}$ is defined in (55) since we have shown that $\boldsymbol{U}^{H}_{\boldsymbol{B}}$ does not destroy the statistical properties of the noise. Finally, we can formulate the spectral form of the WP-MUSIC algorithm.

B. Spectral WP-MUSIC

The spectral WP-MUSIC criterion is

$$\arg\max_{\theta}\mathcal{C}(\theta)^{-1} \quad \text{where}$$

$$\mathcal{C}(\theta) = \left\|\left(\boldsymbol{I}_{L} - \left(\boldsymbol{P}^{\perp}_{\boldsymbol{B}}\hat{\boldsymbol{P}}_{\hat{\boldsymbol{A}}}\right)^{\dagger}\right)\boldsymbol{p}(\theta)\right\|^{2}. \quad (43)$$

The peaks in the pseudospectrum $\mathcal{C}(\theta)^{-1}$ coincide with the unknown DOA. Note that the minimization of $\mathcal{C}(\theta)$ can be interpreted as a generic one-dimensional subspace fitting problem [25].

C. Root WP-MUSIC

The enumerative search procedure associated with the spectral WP-MUSIC criterion is a costly operation. Thanks to the ULA assumption, we expose the "root" version of the WP-MUSIC algorithm which has a lower complexity cost. In addition, it is well known that the "root" version of the MUSIC algorithm is superior to its spectral form [15].

1) Root-CMUSIC Principle: Let $z = e^{-2i\pi\Delta\sin(\theta)}$. The criterion of the root-CMUSIC[4] is given by

$$f_{\text{CMUSIC}}(z) = \boldsymbol{p}\left(\frac{1}{z}\right)^{T}\bar{\boldsymbol{U}}_{\hat{\boldsymbol{Z}}}\bar{\boldsymbol{U}}^{H}_{\hat{\boldsymbol{Z}}}\boldsymbol{p}(z) \quad (44)$$

where $\bar{\boldsymbol{U}}_{\hat{\boldsymbol{Z}}}$ denotes an unitary basis of the noise space $\mathcal{R}(\hat{\boldsymbol{Z}})^{\perp}$ obtained through the methodology introduced in [6] and [11]. Due to the ULA assumption, $\boldsymbol{p}(\cdot)$ has a Vandermonde structure and the DOA estimation problem can be formulated in term of finding the zeros of the above conjugate centrosymmetric polynomial of degree $2L-2$. This symmetry is a consequence of the Hermitian character of projector $\bar{\boldsymbol{U}}_{\hat{\boldsymbol{Z}}}\bar{\boldsymbol{U}}^{H}_{\hat{\boldsymbol{Z}}}$ and the explicit computation of the coefficients of $f_{\text{CMUSIC}}(z)$ denoted by $\{q_{\ell}\}_{\ell\in[1-L:L-1]}$ is given by summing along the diagonal of the projector matrix. In addition, we have $q_{\ell} = q^{*}_{-\ell}$ and q_0 is real and equals to $\text{Tr}(\bar{\boldsymbol{U}}_{\hat{\boldsymbol{Z}}}\bar{\boldsymbol{U}}^{H}_{\hat{\boldsymbol{Z}}}) = L - M$. Moreover, one can easily verify that $f_{\text{CMUSIC}}(z)$ is equal to its reciprocal polynomial [2] and, therefore, if z_m is a zero then z^{*}_{m} is also a zero, i.e., (z_m, z^{*}_{m}) occur in pairs. Note that for the M desired DOA, we have constraint $|z_m| = 1$, i.e., the DOA belong to the unit circle. In presence of noise, the DOA may be extracted (among $2L-2$ possible roots) based on their proximity to the unit circle.

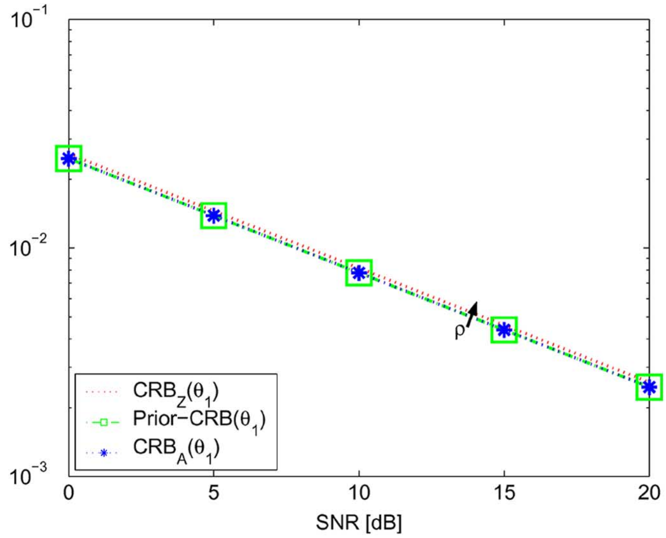
2) Polynomial Form of the Corrective Function and Root WP-MUSIC Algorithm: Here, we follow the same methodology as for the root-CMUSIC approach, and we associate a polynomial form to $f_{\text{COR}}(\theta)$ such as for all unknown DOA, the following polynomial:

$$\begin{aligned}f_{\text{COR}}(z) &= \boldsymbol{p}\left(\frac{1}{z}\right)^{T}\hat{\boldsymbol{E}}^{H}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\hat{\boldsymbol{E}}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\boldsymbol{p}(z) && (45)\\ &= \boldsymbol{p}\left(\frac{1}{z}\right)^{T}\left(\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\boldsymbol{P}_{\boldsymbol{B}}\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\right)^{\dagger}\boldsymbol{p}(z) && (46)\end{aligned}$$

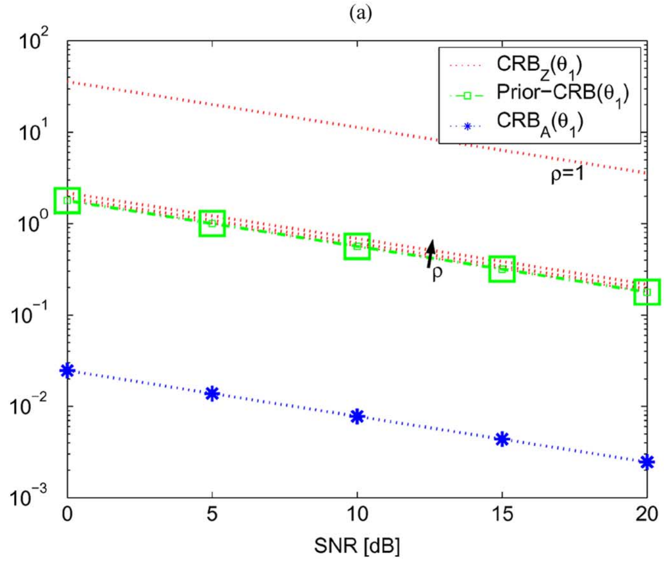
must be zero. By analogy to (42), we have $\hat{\boldsymbol{E}}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})} = (\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\boldsymbol{P}_{\boldsymbol{B}})^{\dagger}$. Then, polynomial (46) is obtained by remarking that

$$\begin{aligned}\hat{\boldsymbol{E}}^{H}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})}\hat{\boldsymbol{E}}_{(\boldsymbol{B}\,\hat{\boldsymbol{A}})} &= \left(\boldsymbol{P}_{\boldsymbol{B}}\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\right)^{\dagger}\left(\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\boldsymbol{P}_{\boldsymbol{B}}\right)^{\dagger} && (47)\\ &= \left(\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\boldsymbol{P}_{\boldsymbol{B}}\boldsymbol{P}_{\boldsymbol{B}}\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\right)^{\dagger} && (48)\\ &= \left(\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\boldsymbol{P}_{\boldsymbol{B}}\hat{\boldsymbol{P}}^{\perp}_{\hat{\boldsymbol{A}}}\right)^{\dagger}. && (49)\end{aligned}$$

[4]Here again, we expose the "root" version of the WP-MUSIC algorithm based on (28) since we consider that this expression clearly highlights the link between the root WP-MUSIC and the root-CMUSIC. However, for the two reasons explained in Section V, it is preferable to use (27) to really implement the root WP-MUSIC based on the resolution of polynomial $\boldsymbol{p}(1/z)^{T}(\boldsymbol{I}_{L} - \hat{\boldsymbol{E}}^{H}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})})(\boldsymbol{I}_{L} - \hat{\boldsymbol{E}}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})})\boldsymbol{p}(z)$ where projector $\hat{\boldsymbol{E}}_{(\hat{\boldsymbol{A}}\,\boldsymbol{B})}$ is given by (42).



(a)



(b)

Fig. 1. CRB versus SNR [dB]. (a) Widely spaced DOA ($\theta = [5^\circ\ 80^\circ]$). (b) Closely spaced DOA ($\theta = [5^\circ\ 5.2^\circ]$).

Note that due to the fact that $\hat{\boldsymbol{E}}^H_{(\boldsymbol{B}\ \hat{\boldsymbol{A}})}\hat{\boldsymbol{E}}_{(\boldsymbol{B}\ \hat{\boldsymbol{A}})}$ is Hermitian, the coefficients of $f_{\mathrm{COR}}(z)$, noted $\{p_\ell\}_{\ell\in[1-L:L-1]}$, are conjugate centrosymmetry, i.e., $p_\ell = p^*_{-\ell}$, $p_0 = \mathrm{Tr}(\hat{\boldsymbol{E}}^H_{(\boldsymbol{B}\ \hat{\boldsymbol{A}})}\hat{\boldsymbol{E}}_{(\boldsymbol{B}\ \hat{\boldsymbol{A}})})$ and therefore (z_m, z^*_m) occur in pairs. Consequently, the root WP-MUSIC is based on the following result.

Theorem 3: The S roots of polynomial $\mathcal{C}(z) = f_{\mathrm{CMUSIC}}(z) + f_{\mathrm{COR}}(z)$ where $f_{\mathrm{CMUSIC}}(z)$ and $f_{\mathrm{COR}}(z)$ has been defined in (44) and (45), respectively, are the set of the DOA without the subset of the known DOA.

Proof: See Appendix E.

VI. Numerical Simulations

A. Numerical Analysis of the Prior-CRB

The geometry of the array is Uniform and Linear (ULA) of $L = 10$ sensors and $T = 100$ snapshots. So, the Rayleigh resolution is about 0.49. To illustrate the comparison between the three derived bounds: the $\mathrm{CRB}_{\boldsymbol{A}}$, the $\mathrm{CRB}_{\boldsymbol{Z}}$ and the $\mathrm{P-CRB}$, we consider two situations.

1) One Known DOA and One Unknown: In this situation, θ_2 is known and θ_1 have to be estimated. The covariance matrix of the sources is given by $\boldsymbol{R}_{\boldsymbol{\Lambda}} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{\Lambda}_{\hat{\boldsymbol{A}}}} & \rho \\ \rho^* & \boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}} \end{bmatrix}$ with $\boldsymbol{R}_{\boldsymbol{\Lambda}_{\hat{\boldsymbol{A}}}} = \boldsymbol{R}_{\boldsymbol{\Lambda}_{\mathbf{B}}} = 1$ and $|\rho|^2 \in [0 : 0.99]$. The coherent scenario is considered for singular spatial covariance, i.e., for $\rho = 1$. In Fig. 1(a), we consider uncorrelated, correlated and coherent sources with widely spaced DOA. In that case, the Prior-CRB is comparable with the CRB for a single or two DOA. So, in this situation, the knowledge of θ_2 cannot help the estimation of θ_1. This situation confirms property (vi) and (iii) even if the sources are correlated.

In Fig. 1(b), the DOA are closely spaced. In this case, the CRB for one DOA is much lower than the P-CRB and the CRB for two DOA, as expected in (iv). In addition, for block-diagonal source covariance, the P-CRB and the CRB for two DOA are merged according to property (iii). This illustrates in particular property (iv). The important point is that even if the sources are highly correlated, the gain associated with the P-CRB with respect to the CRB for two DOA is small. Inversely, for coherent sources with closely spaced DOA, the P-CRB is much lower than the CRB for two DOA. This observation illustrates property (v). In that case, prior knowledge is beneficial.

Finally, in Fig. 2, we vary the number of sensors for closely spaced DOA. We can see that the CRB are asymptotically merged which confirms property (ii).

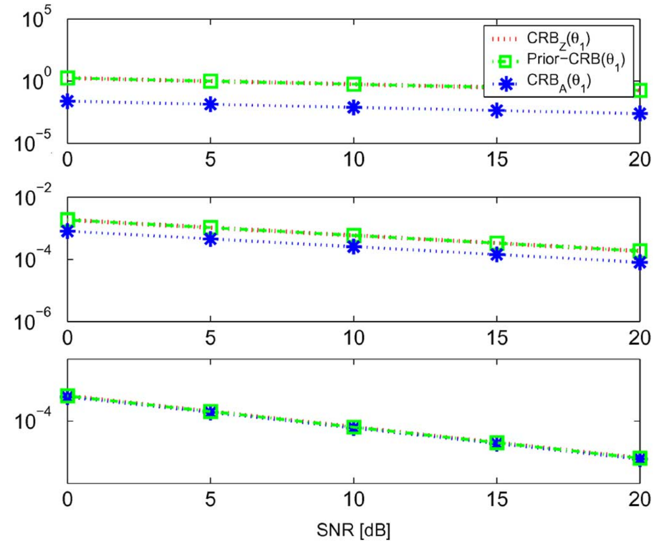


Fig. 2. CRB versus SNR [dB], Closely spaced DOA, (top) $L = 10$, (middle) $L = 300$, (bottom) $L = 1000$.

2) Two Known DOA and One Unknown: Here, we know θ_2 and θ_3 and we want to estimate θ_1. The spatial covariance is given by

$$\boldsymbol{R}_{\boldsymbol{\Lambda}} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{\Lambda}_{\hat{\boldsymbol{A}}}} & \boldsymbol{v} \\ \boldsymbol{v}^H & \boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}} \end{bmatrix} \tag{50}$$

where $\boldsymbol{v} = [\rho_{12}\ \rho_{13}]$, $\boldsymbol{R}_{\boldsymbol{\Lambda}_{\hat{\boldsymbol{A}}}} = 1$ and

$$\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}} = \begin{bmatrix} 1 & \rho_{23} \\ \rho^*_{23} & 1 \end{bmatrix} \tag{51}$$

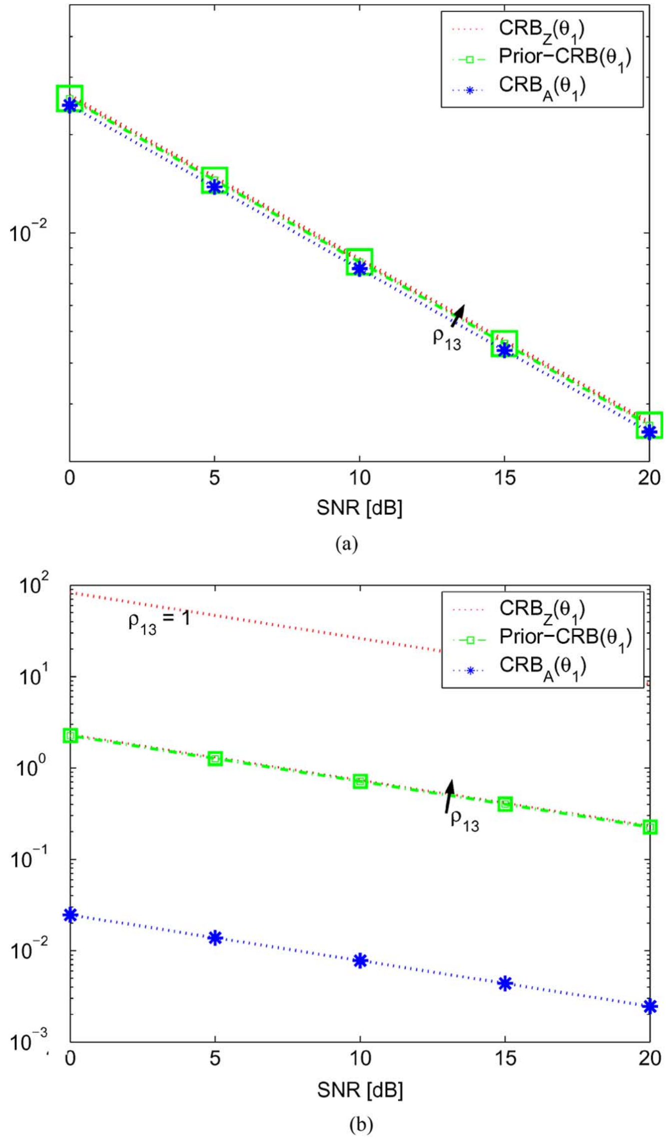


Fig. 3. CRB versus SNR [dB]. (a) Widely spaced DOA ($\theta = [5^\circ\ 80^\circ 40^\circ]$). (b) Closely spaced DOA ($\theta = [5^\circ\ 5.2^\circ\ 7^\circ]$).

with $|\rho_{12}|^2 = 0.32$, $|\rho_{23}|^2 = 0.99$, and ρ_{13} varies according to ρ in the previous section. So, the known DOA are highly correlated and the correlation coefficient between the unknown DOA associated with the first and third sources varies until to the coherent scenario. On Fig. 3, we have drawn the CRB for closely spaced and for widely spaced DOA. Like in the previous situation, the P-CRB indicates that the exploitation of prior knowledge is interesting only for coherent sources with closely spaced DOA.

B. Illustration of the WP-MUSIC Algorithm

Here, we consider a numerical example to illustrate the WP-MUSIC algorithm. On Fig. 4(a), we have drawn the pseudospectrums of the CMUSIC and the WP-MUSIC algorithms for three DOA where one is known and two others have to be estimated. First, note on the Prior-MUSIC pseudospectrum that the known DOA at 100° has been efficiently cancelled from the CMUSIC pseudospectrum without altering the unknown

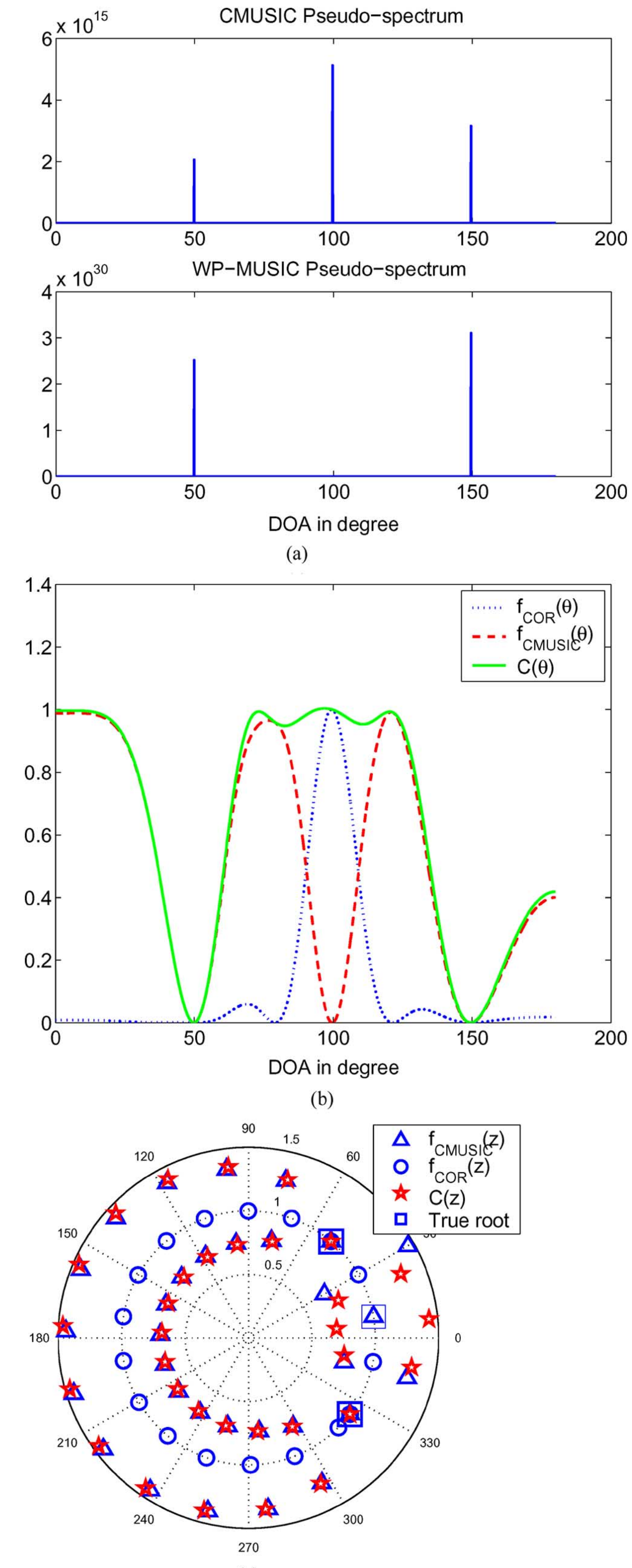


Fig. 4. (a) CMUSIC and WP-MUSIC pseudospectrums for three DOA (one known and two unknown). (b) $f_{\mathrm{CMUSIC}}(\theta), f_{\mathrm{COR}}(\theta)$ and $\mathcal{C}(\theta)$ for $L = 18$ sensors. (c) Zero location with respect to the unit circle.

one. In contrast with the CMUSIC algorithm, we can note on Fig. 4(b) that $\mathcal{C}(\theta)$ has only two null values at 50° and 150°.

In Fig. 4(c), we have drawn the zero location with respect to the unit circle for the root-CMUSIC and root WP-MUSIC algorithms. Note that the zeros occur in pairs, as expected. However, in presence of noise, selecting the zeros (with unit modulus constraint) based only on $f_{\mathrm{COR}}(z)$ is a difficult task due to their proximity to the unit circle. So, a decision only based on $f_{\mathrm{COR}}(z)$ seems ineffective. Inversely, note that a decision on criterion $\mathcal{C}(z)$ is a more practicable task.

C. Performances of the Algorithms

1) Accuracy of the Proposed Methods: We assume that the source covariance is block-diagonal. The tested methods are

- WP-MUSIC: The MUSIC algorithm with prior knowledge based on the obliquely weighted pseudoinverse (*cf.* Section IV-B).
- P-MUSIC: The MUSIC algorithm with prior knowledge based on standard LS resolution (*cf.* Section IV-A).
- P-MUSIC (SI): The MUSIC algorithm with prior knowledge based on the implementation of the oblique projector proposed by McCloud and Scharf [12].
- MUSIC: The standard root version of the MUSIC algorithm.
- CMUSIC: The root version of the constrained MUSIC algorithm presented by DeGroat *et al.* [6].

The accuracy of the DOA of interest estimation is measured through the Standard Deviation (Std) which is defined as the root of the MSE. Each simulation is based on 1000 Monte Carlo trials. In Fig. 5(a), we consider widely spaced DOA, e.g., $\theta = [80^\circ\ 5^\circ]$. In this situation, all the tested algorithms are equivalent. On Fig. 5(b), we choose closely spaced DOA, e.g., $\theta = [8^\circ\ 5^\circ]$ which corresponds to a distance inter-DOA much lower than the Rayleigh resolution for 10 sensors. In this scenario, the CRB for two DOA and the P-CRB for the DOA of interest are merged, as expected in property (iii) of Theorem 1 (cf. Section III-D) in the context of more than two DOA. We can note that for sufficient SNR, the CMUSIC and the MUSIC algorithms reach these bounds but they cannot outperform it for closely spaced DOA and for block-diagonal source covariance. The WP-MUSIC, P-MUSIC, and P-MUSIC (SI) show Std close to the CRB for only one DOA at high SNR. By the light of this example, we can say that most of the influence of the known DOA has been efficiently cancelled by the proposed algorithms. This is not the case for the CMUSIC algorithm. According to Fig. 5(b) and (c), the WP-MUSIC algorithm is slightly more efficient than the P-MUSIC algorithm at low SNR ($\approx$10 dB) and for closely spaced DOA. This observation confirms the discussion in Sections IV-A and IV.B. Finally, we perform in Fig. 5(d), (e), and (f), some experiments with two known DOA and one unknown. The conclusions are similar to the more simple case of one unknown and one known.

2) Robustness to a Small Error on the a Priori: The scenario is the same as for Fig. 5(a), i.e., the known DOA is $\theta_1 = 80^\circ$ and the unknown DOA is $\theta_2 = 5^\circ$. We perturb θ_2 according to $\pm\delta_{\theta_2} = 5^\circ$ and we compute the standard deviation for the DOA of interest, θ_1. This scenario is repeated for different SNR and number of sensors. The number of snapshots is equal to 100.

Fig. 6(a) shows that without noise, the P-MUSIC (SI) and the CMUSIC are very sensitive to a small error on the known DOA. Inversely, the P-MUSIC and WP-MUSIC algorithms are more robust. These observations are confirmed by Fig. 6(b) where the SNR is equal to 30 dB. Indeed, we can note the remarkable robustness of the P-MUSIC and WP-MUSIC algorithms since the Std associated with these methods follow a flat curve. This is a clear advantage of these approaches. In Fig. 6(c), we can see that all the tested methods have similar robustness at low SNR (0 dB). In this situation, the error induced by the noise dominates the error associated with the error on the known DOA.

In Fig. 6(d), we have drawn the Std for the tested methods without noise and for a large number of sensor $(L = 100)$. As expected, in this asymptotic regime, the CMUSIC, the P-MUSIC and the WP-MUSIC have the same efficiency and robustness. Note that in this case, the bad results for the P-MUSIC (SI) algorithm. These observations are confirmed in Fig. 6(e) and (f) in the noisy situation.

VII. Conclusion of the Simulation Part

1) For a small number of sensors, the WP-MUSIC shows the best accuracy and the best robustness to a small error on the known DOA. In particular, the accuracy of this algorithm is near the CRB associated with the subspace of interest. Consequently, the influence of the known DOA is almost cancelled.
2) As explained in Sections IV-A, IV-B and in the simulation part, the P-MUSIC algorithm is slightly less efficient than the WP-MUSIC algorithm but its robustness is comparable. This is a valuable solution.
3) The P-MUSIC (SI) shows a comparable accuracy as the P-MUSIC and the WP-MUSIC algorithms at high SNR but this algorithm is less efficient in severely noisy situations. In addition, its robustness is weak. So, essentially for the latest reason mentioned, we prefer the implementation of the oblique projector introduced in Section V-A rather than the one presented by McCloud and Scharf.[5]
4) For a small number of sensors and for closely spaced DOA associated with block-diagonal source covariance, the P-MUSIC, the WP-MUSIC, and the P-MUSIC (SI) algorithms outperform the CMUSIC algorithms, in particular at high SNR where the CMUSIC algorithm is lower bounded by the CRB over the whole space. For a large number of sensors and/or widely spaced sources, this algorithm is equivalent to the ones based on oblique projectors.

VIII. Conclusion

In this paper, we have presented a subspace-based solution to estimate S DOA among M using the knowledge of $M - S$ known DOA. In a first part of this paper, we have derived and analyzed the CRB associated with the orthogonal deflation of the signal subspace and we have shown several limitations of this approach. Consequently in the second part, we have proposed alternative solutions based on an oblique deflation of the

[5]Note that in this work, the proposed methodology also assumes the block-diagonal structure of the spatial covariance of the sources.

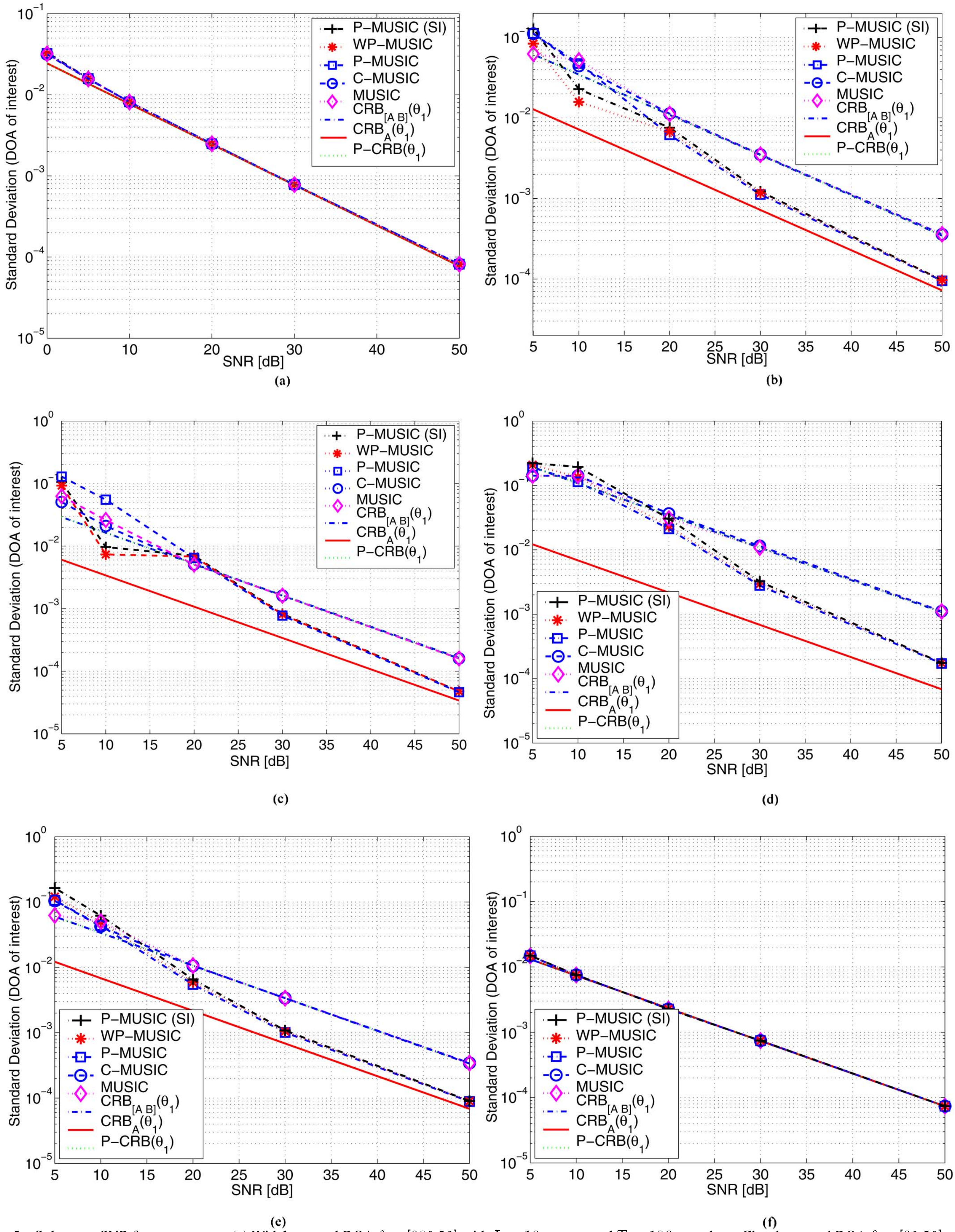


Fig. 5. Std versus SNR for two sources. (a) Widely spaced DOA $\theta = [80^\circ\ 5^\circ]$ with $L = 10$ sensors and $T = 100$ snapshots. Closely spaced DOA $\theta = [8^\circ\ 5^\circ]$ with $L = 10$ sensors. (b) With $T = 100$ snapshots. (c) With $T = 500$ snapshots. Std versus SNR for three sources. (d) Closely spaced DOA $\theta = [8^\circ\ 5^\circ\ 12^\circ]$ with $L = 10$ sensors and $T = 100$ snapshots. (e) Closely spaced DOA $\theta = [8^\circ\ 5^\circ\ 80^\circ]$ $L = 10$ sensors and $T = 100$ snapshots. (f) Widely spaced DOA $\theta = [8^\circ\ 50^\circ\ 80^\circ]$ $L = 10$ sensors and $T = 100$ snapshots.

signal subspace. We show that the proposed algorithm, called Prior-MUSIC, mitigates almost the influence of the known DOA on the DOA of interest in particular when the DOA are closely spaced and the source covariance is block-diagonal. Finally, the oblique projector framework provides a suitable way to integrate prior knowledge into subspace-based methods and more generally into subspace fitting techniques.

APPENDIX

A. Brief Discussion of Oblique Projectors

This appendix is dedicated to oblique projections [1]. In particular, we recall that the only requirement of a matrix $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$ to be a projector is $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$ is idempotent, i.e., $\boldsymbol{E}^2_{(\boldsymbol{X}\,\boldsymbol{Y})} = \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$. Let $\mathcal{R}(\boldsymbol{X})$ and $\mathcal{R}(\boldsymbol{Y})$ be subspaces of $\mathbb{C}^L$ that intersect trivially, i.e., $\mathcal{R}(\boldsymbol{X}) \cap \mathcal{R}(\boldsymbol{Y}) = \{\boldsymbol{0}\}$. Then, the projector on $\mathcal{R}(\boldsymbol{X})$ along $\mathcal{R}(\boldsymbol{Y})$ is the linear operator $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$ satisfying:

- $\forall \boldsymbol{x} \in \mathcal{R}(\boldsymbol{X}),\ \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}\boldsymbol{x} = \boldsymbol{x}$;
- $\forall \boldsymbol{y} \in \mathcal{R}(\boldsymbol{Y}),\ \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}\boldsymbol{y} = \boldsymbol{0}$;
- $\forall \boldsymbol{z} \in \mathbb{C}^L,\ \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}\boldsymbol{z} \in \mathcal{R}(\boldsymbol{X})$.

The geometric interpretation of the above properties is $\forall \boldsymbol{s} = \boldsymbol{x}+\boldsymbol{y}+\boldsymbol{z} \in \mathbb{C}^L$ where $\boldsymbol{x} \in \mathcal{R}(\boldsymbol{X})$, $\boldsymbol{y} \in \mathcal{R}(\boldsymbol{Y})$ and $\boldsymbol{z} \in (\mathcal{R}(\boldsymbol{X}) \cup \mathcal{R}(\boldsymbol{Y}))^\perp$ then $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}\boldsymbol{s} = \boldsymbol{x}$. So, the complex Euclidean Space is decomposed according to $\mathbb{C}^L = (\mathcal{R}(\boldsymbol{X}) \cup \mathcal{R}(\boldsymbol{Y}))^\perp \oplus \mathcal{R}(\boldsymbol{X}) \oplus \mathcal{R}(\boldsymbol{Y})$. Let V be a complex matrix having full column rank, obtaining by the concatenation of matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ according to $\boldsymbol{V} = [\boldsymbol{X}\ \boldsymbol{Y}]$. The orthogonal projector onto $\mathcal{R}(\boldsymbol{V})$ is then defined as $\boldsymbol{P}_{\boldsymbol{V}} = \boldsymbol{V}\boldsymbol{V}^\dagger = \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})} + \boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})}$ and

$$\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})} = \boldsymbol{X}\left(\boldsymbol{X}^H \boldsymbol{P}^\perp_{\boldsymbol{Y}} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^H \boldsymbol{P}^\perp_{\boldsymbol{Y}}. \tag{52}$$

This property is important since it highlights the link between the orthogonal projector $\boldsymbol{P}_{\boldsymbol{V}}$ and oblique projectors $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$ and $\boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})}$. In addition, the ranges for $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$ and $\boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})}$ are $\mathcal{R}(\boldsymbol{X})$ and $\mathcal{R}(\boldsymbol{Y})$, respectively, and the null spaces for $\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}$ and $\boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})}$ are $\mathcal{R}(\boldsymbol{Y}) \oplus \mathcal{R}(\boldsymbol{V}^\perp)$ and $\mathcal{R}(\boldsymbol{X}) \oplus \mathcal{R}(\boldsymbol{V}^\perp)$, respectively. A useful rewritten of $\boldsymbol{P}^\perp_{\boldsymbol{V}}$ can be deduced from (VIII-A) according to

$$\boldsymbol{P}^\perp_{\boldsymbol{V}} = \boldsymbol{I}_L - \boldsymbol{P}_{\boldsymbol{V}} = \boldsymbol{I}_L - \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})} - \boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})}. \tag{53}$$

Finally, note that

$$\boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})}\boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})} = \boldsymbol{E}_{(\boldsymbol{X}\,\boldsymbol{Y})}\boldsymbol{E}_{(\boldsymbol{Y}\,\boldsymbol{X})} = \boldsymbol{0}. \tag{54}$$

B. Demonstration of Proposition 1

First, note that an ordered eigendecomposition of any rank-$(L-M+S)$ idempotent matrices is

$$\begin{aligned}\boldsymbol{P}^\perp_{\boldsymbol{B}} &= [\boldsymbol{U}_{\boldsymbol{B}} \quad \times] \begin{bmatrix} \boldsymbol{I}_{L-M+S} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} [\boldsymbol{U}_{\boldsymbol{B}} \quad \times]^H \\ &= \boldsymbol{U}_{\boldsymbol{B}} \boldsymbol{U}^H_{\boldsymbol{B}} \end{aligned} \tag{55}$$

where $\boldsymbol{U}_{\boldsymbol{B}}$ is the $L-M+S$ first columns of the left eigenfactor. Thanks to the property that the noise is zero-mean and $\boldsymbol{U}^H_{\boldsymbol{B}}\boldsymbol{U}_{\boldsymbol{B}} = \boldsymbol{I}_{L-M+S}$, it is not difficult to see that

$$\begin{aligned}\mathbb{E}\left(\tilde{\boldsymbol{y}}_{(\text{prior})}\right) &= \left(\boldsymbol{I}_T \otimes \boldsymbol{U}^H_{\boldsymbol{B}}\right)\left(\boldsymbol{I}_T \otimes \left(\boldsymbol{U}_{\boldsymbol{B}}\boldsymbol{U}^H_{\boldsymbol{B}}\right)\right) \\ &\quad \times (\boldsymbol{I}_T \otimes \boldsymbol{Z})\boldsymbol{\lambda} \end{aligned} \tag{56}$$

$$\begin{aligned} &= \left(\boldsymbol{I}_T \otimes \left(\boldsymbol{U}^H_{\boldsymbol{B}}\boldsymbol{A}\right)\right)\bar{\boldsymbol{\lambda}} && (57)\\ &= \tilde{\boldsymbol{x}}_{(\text{prior})}. && (58)\end{aligned}$$

The error covariance, noted $\tilde{\boldsymbol{\Gamma}}$, is given by the covariance of the following centered signal $\tilde{\boldsymbol{y}}_{(\text{prior})} - \tilde{\boldsymbol{x}}_{(\text{prior})} = (\boldsymbol{I}_T \otimes \boldsymbol{U}^H_{\boldsymbol{B}})\boldsymbol{n}$ then

$$\begin{aligned}\tilde{\boldsymbol{\Gamma}} &= \mathbb{E}\left(\left(\tilde{\boldsymbol{y}}_{(\text{prior})} - \tilde{\boldsymbol{x}}_{(\text{prior})}\right)\right. \\ &\qquad \left. \times \left(\tilde{\boldsymbol{y}}_{(\text{prior})} - \tilde{\boldsymbol{x}}_{(\text{prior})}\right)^H\right) && (59)\\ &= \left(\boldsymbol{I}_T \otimes \boldsymbol{U}^H_{\boldsymbol{B}}\right)(\sigma^2 \boldsymbol{I}_{LT})(\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}) && (60)\\ &= \sigma^2 \boldsymbol{I}_{LT}. && (61)\end{aligned}$$

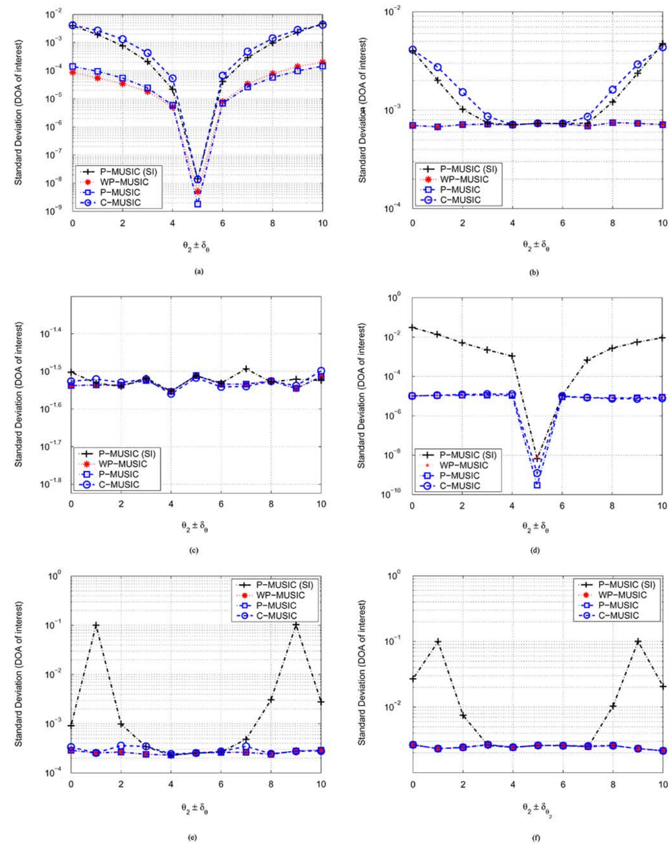


Fig. 6. Std versus Error on the known DOA with $L = 10$ sensors. (a) Without noise. (b) With $\mathrm{SNR} = 30$ dB. (c) With $\mathrm{SNR} = 0$ dB. $L = 100$ sensors and $T = 100$ snapshots. (d) Without noise. (e) With $L = 30$ sensors and $T = 100$ snapshots and $\mathrm{SNR} = 30$ dB. (f) With $\mathrm{SNR} = 0$ dB.

C. Demonstration of Theorem 1

We recall that the signal plus nuisance model parameter vector by $\chi = [\chi'^T\ \sigma^2]^T$ where $\chi' = [\boldsymbol{\theta}^{(S)^T}\ \bar{\boldsymbol{\lambda}}^T_{\mathcal{R}}\ \bar{\boldsymbol{\lambda}}^T_{\mathcal{I}}]^T$, $\boldsymbol{\theta}^{(S)} = [\theta_1 \ldots \theta_S]^T$, $\bar{\boldsymbol{\lambda}}_{\mathcal{R}} = \Re e\{\bar{\boldsymbol{\lambda}}\}$ and $\bar{\boldsymbol{\lambda}}_{\mathcal{I}} = \Im m\{\bar{\boldsymbol{\lambda}}\}$. The first term in (14) is associated with the noise and is given by

$$\mathrm{Tr}\left(\tilde{\boldsymbol{\Gamma}}^{-1}\frac{\partial\tilde{\boldsymbol{\Gamma}}}{\partial\boldsymbol{\chi}}\tilde{\boldsymbol{\Gamma}}^{-1}\frac{\partial\tilde{\boldsymbol{\Gamma}}}{\partial\boldsymbol{\chi}}\right) = \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{(L-M+S)T}{\sigma^4}\end{bmatrix}. \tag{62}$$

The second term which involved the partial derivatives of the noise-free model with respect to the parameter vector $\boldsymbol{\chi}$ can be expressed according to

$$\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \bar{\boldsymbol{\lambda}}_{\mathcal{R}}} = \boldsymbol{I}_T \otimes \left(\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}\right) \tag{63}$$

$$\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \bar{\boldsymbol{\lambda}}_{\mathcal{I}}} = i\boldsymbol{I}_T \otimes \left(\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}\right) \tag{64}$$

$$\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \boldsymbol{\theta}^{(S)}} = \left(\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}^H\right) \boldsymbol{C} \tag{65}$$

where

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{p}'(\theta_1)\alpha_1(1) & \cdots & \boldsymbol{p}'(\theta_S)\alpha_S(1) \\ \vdots & & \vdots \\ \boldsymbol{p}'(\theta_1)\alpha_1(T) & \cdots & \boldsymbol{p}'(\theta_S)\alpha_S(T) \end{bmatrix} \tag{66}$$

with $\boldsymbol{p}'(\theta_k) = \frac{\partial \boldsymbol{p}(\theta)}{\partial \theta}|_{\theta=\theta_k}$. So, we obtain

$$\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \boldsymbol{\chi}} = \left[\, \left(\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}^H\right) \boldsymbol{C} \quad \boldsymbol{I}_T \otimes \left(\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}\right) \quad i\boldsymbol{I}_T \otimes \left(\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}\right) \quad \boldsymbol{0} \,\right].$$

The P-CRB is given by the matrix at the bottom of the page.

As the nuisance and signal parameters are decoupled, the P-CRB has a block-diagonal structure. To obtain the P-CRB for subvector $\boldsymbol{\theta}^{(S)}$, we follow the block-diagonalization method introduced in [18, p. 390]. Finally, we have

$$\text{P}-\text{CRB}(\boldsymbol{\theta}^{(S)}) = \frac{\sigma^2}{2} \left[\Re e \left\{ \boldsymbol{C}^H (\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}) \boldsymbol{P}^{\perp}_{\boldsymbol{I}_T \otimes (\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A})} (\boldsymbol{I}_T \otimes \boldsymbol{U}_{\boldsymbol{B}}^H) C \right\} \right]^{-1}$$

where $\boldsymbol{P}^{\perp}_{\boldsymbol{I}_T \otimes (\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A})} = \boldsymbol{I}_{(L-M+S)T} - (\boldsymbol{I}_T \otimes (\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}))(\boldsymbol{I}_T \otimes (\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}))^{\dagger}$.

After some straightforward algebraic derivations and for sufficient number of snapshots, it comes

$$\text{P}-\text{CRB}(\boldsymbol{\theta}^{(S)}) = \frac{\sigma^2}{2T} \left[\Re e \left\{ \left[\boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{U}_{\boldsymbol{B}} \boldsymbol{P}^{\perp}_{\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}} \boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{D}_{\boldsymbol{A}} \right] \bullet \boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}} \right\} \right]^{-1}.$$

Let $\boldsymbol{E}_{(\boldsymbol{B}\ \boldsymbol{A})}$ (respectively, $\boldsymbol{E}_{(\boldsymbol{A}\ \boldsymbol{B})}$) be the oblique projector on $\mathcal{R}(\boldsymbol{B})$ (respectively, $\mathcal{R}(\boldsymbol{A})$) along $\mathcal{R}(\boldsymbol{A})$ (respectively, $\mathcal{R}(\boldsymbol{B})$) defined in (52). Based on these operators, we have the following property:

$$\boldsymbol{U}_{\boldsymbol{B}} \boldsymbol{P}^{\perp}_{\boldsymbol{U}_{\boldsymbol{B}}^H \boldsymbol{A}} \boldsymbol{U}_{\boldsymbol{B}}^H = \boldsymbol{P}_{\boldsymbol{B}}^{\perp} (\boldsymbol{I}_L - \underbrace{\boldsymbol{A}(\boldsymbol{A}^H \boldsymbol{P}_{\boldsymbol{B}}^{\perp} \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{P}_{\boldsymbol{B}}^{\perp}}_{\boldsymbol{E}_{(\boldsymbol{A}\ \boldsymbol{B})}}) \tag{67}$$

$$= \boldsymbol{P}_{\boldsymbol{B}}^{\perp} \left(\boldsymbol{P}_{\boldsymbol{Z}}^{\perp} + \boldsymbol{E}_{(\boldsymbol{B}\ \boldsymbol{A})} \right) \tag{68}$$

$$= \boldsymbol{P}_{\boldsymbol{B}}^{\perp} \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \tag{69}$$

$$= \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \tag{70}$$

where to obtain (68) [respectively, (69)], we use (53) [respectively, (54)]. Consequently, the P-CRB can be simplified to (15).

D. Demonstration of Theorem 2

To show property (i), it is equivalent to prove that $\text{P}-\text{CRB}(\boldsymbol{\theta}^{(S)}) - \text{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$ is a positive semidefinite (psd) matrix. First, using $\boldsymbol{P}_{\boldsymbol{A}}^{\perp} \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} = \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \boldsymbol{P}_{\boldsymbol{A}}^{\perp} = \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}$, it is straightforward to see that $\boldsymbol{P}_{\boldsymbol{A}}^{\perp} - \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}$ is idempotent. Therefore, the eigenvalues of $\boldsymbol{P}_{\boldsymbol{A}}^{\perp} - \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}$ is 1 or 0 and, thus, $\boldsymbol{P}_{\boldsymbol{A}}^{\perp} - \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}$ is psd. Next, as $\boldsymbol{D}_{\boldsymbol{A}}$ is a nondeficient matrix, $\boldsymbol{D}_{\boldsymbol{A}}^H (\boldsymbol{P}_{\boldsymbol{A}}^{\perp} - \boldsymbol{P}_{\boldsymbol{Z}}^{\perp}) \boldsymbol{D}_{\boldsymbol{A}}$ is also psd. Finally using that (1) the Hadamard product of two psd matrices is also a psd matrix (*cf.* result R19 in [18]) and (2) the real part of a psd matrix is psd itself, we prove (i).

Property (ii) can be proved in the following manner. In [21], it has been shown that $\text{CRB}_{\boldsymbol{A}}(\theta_m), \text{CRB}_{\boldsymbol{Z}}(\theta_m) \xrightarrow{L \gg 1} \frac{6}{L^3 T} \frac{\sigma^2}{[\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}}]_{mm}}$. For the P-CRB, we have $\frac{1}{L^3} \boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{D}_{\boldsymbol{A}} \xrightarrow{L \gg 1} \frac{1}{3} \boldsymbol{I}_S$, $\frac{1}{L^2} \boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{Z} \xrightarrow{L \gg 1} \frac{i}{2} [\boldsymbol{I}_S \quad 0]_{S \times M}$ and $\frac{1}{L} \boldsymbol{Z}^H \boldsymbol{Z} \xrightarrow{L \gg 1} \boldsymbol{I}_M$ then

$$\boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \boldsymbol{D}_{\boldsymbol{A}} \xrightarrow{L \gg 1} \frac{L^3}{3} \boldsymbol{I}_S - \frac{L^3}{4} [\boldsymbol{I}_S \quad 0] \boldsymbol{I}_M \begin{bmatrix} \boldsymbol{I}_S \\ 0 \end{bmatrix} = \frac{L^3}{12} \boldsymbol{I}_S. \tag{71}$$

Consequently, using the definition of the P-CRB, it comes

$$\text{P}-\text{CRB}(\theta_m) \xrightarrow{L \gg 1} \frac{6}{L^3 T} \frac{\sigma^2}{[\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}}]_{mm}}$$

which proves property (ii).

Next, note that the CRB over the whole space can be rewritten according to

$$\text{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(M)}) = \frac{\sigma^2}{2T} \left[\Re e \left\{ \begin{bmatrix} \boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \boldsymbol{D}_{\boldsymbol{A}} & \boldsymbol{D}_{\boldsymbol{A}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \boldsymbol{D}_{\boldsymbol{B}} \\ \boldsymbol{D}_{\boldsymbol{B}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \boldsymbol{D}_{\boldsymbol{A}} & \boldsymbol{D}_{\boldsymbol{B}}^H \boldsymbol{P}_{\boldsymbol{Z}}^{\perp} \boldsymbol{D}_{\boldsymbol{B}} \end{bmatrix} \bullet \boldsymbol{R}_{\boldsymbol{\Lambda}} \right\} \right]^{-1} \tag{73}$$

where $\boldsymbol{D}_{\boldsymbol{B}} = [\boldsymbol{p}'(\theta_{S+1}) \ldots \boldsymbol{p}'(\theta_M)]$. In addition, suppose that $\boldsymbol{R}_{\boldsymbol{\Lambda}}$ is block-diagonal in (73) then it is easy to deduce property (iii) since

$$\text{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(M)}) = \begin{bmatrix} \text{P}-\text{CRB}(\boldsymbol{\theta}^{(S)}) & \boldsymbol{0} \\ \boldsymbol{0} & \times \end{bmatrix}. \tag{74}$$

Note that this relation holds for widely or closely spaced DOA.

If the DOA are widely spaced, the influence between the DOA is weak then it is well known that $\text{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(S)}) \approx \text{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$. Consequently, according to property (iii), the $\text{P}-\text{CRB}(\boldsymbol{\theta}^{(S)})$ reaches its minimum near the $\text{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$ by superior values [*cf.* relation (i)], which proves property (vi). Inversely, for closely spaced DOA, the $\text{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$ is invariant

$$\text{P}-\text{CRB}(\boldsymbol{\chi}) = \begin{bmatrix} \frac{\sigma^2}{2} \left[\Re e \left\{ \left(\frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \boldsymbol{\chi}'} \right)^H \frac{\partial \tilde{\boldsymbol{x}}_{(\text{prior})}}{\partial \boldsymbol{\chi}'} \right\} \right]^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{\sigma^4}{(L-M+S)T} \end{bmatrix}.$$

while the inverse of $\mathrm{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(S)})$ is near the singularity (large condition number with respect to the inversion). Thus, $\mathrm{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(S)}) \gg \mathrm{CRB}_{\boldsymbol{A}}(\boldsymbol{\theta}^{(S)})$. This fact together with property (iii) give property (iv).

Now suppose that $\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}}$ is deficient (coherent sources), then due to its block structure, $\boldsymbol{R}_{\boldsymbol{\Lambda}}$ is also deficient.[6] In addition, if we consider closely spaced DOA then $\mathrm{CRB}_{\boldsymbol{Z}}(\boldsymbol{\theta}^{(S)})$ takes a large value as the inverse of the Hadamard product of two (near) singular matrices. In the same time and through projector $\boldsymbol{P}_{\boldsymbol{Z}}^{\perp}$, the P-CRB remains sensitive to the known DOA but insensitive to the correlation between the sources associated with the known DOA, i.e., to the spatial covariance $\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}}$. Consequently, property (v) holds. For correlated sources, the problem is more complicated and we have deferred the discussion to the simulation part.

E. Demonstration of Theorem 3

As we know (z_m, z_m^*) occur in pairs, we can give the factorized forms of polynomials $f_{\mathrm{CMUSIC}}(z)$ and $f_{\mathrm{COR}}(z)$ according to

$$\begin{aligned} f_{\mathrm{CMUSIC}}(z) &= \prod_{m=1}^{S}(z-z_m)\left(z-\frac{1}{z_m{}^*}\right)^* \prod_{m=S+1}^{M}(z-z_m) \\ &\times \left(z-\frac{1}{z_m{}^*}\right)^* \prod_{m=1}^{L-M-1}(z-z'_m)\left(z-\frac{1}{z'^{\,*}_m}\right)^* \end{aligned} \tag{75}$$

and

$$\begin{aligned} f_{\mathrm{COR}}(z) &= \prod_{m=1}^{S}(z-z_m)\left(z-\frac{1}{z_m{}^*}\right)^* \\ &\times \prod_{m=1}^{L-S-1}(z-z''_m)\left(z-\frac{1}{z''^{\,*}_m}\right)^* \end{aligned} \tag{76}$$

where $\{z_m\}$ are the desired (known or unknown) zeros and $\{z'_m\}$ and $\{z''_m\}$ are the extraneous zeros. Based on (75) and (76), $\mathcal{C}(z)$ admits the following factorization:

$$\mathcal{C}(z) = I(z)Q(z) \tag{77}$$

where

$$I(z) = \prod_{m=1}^{S}(z-z_m)\left(z-\frac{1}{z_m{}^*}\right)^*$$

and

$$\begin{aligned} Q(z) = &\prod_{m=S+1}^{M}(z-z_m)\left(z-\frac{1}{z_m{}^*}\right)^* \prod_{m=1}^{L-M-1}(z-z'_m) \\ &\times \left(z-\frac{1}{z'^{\,*}_m}\right)^* + \prod_{m=1}^{L-S-1}(z-z''_m)\left(z-\frac{1}{z''^{\,*}_m}\right)^*. \end{aligned}$$

[6] We have $\det(\boldsymbol{R}_{\boldsymbol{\Lambda}}) = \det(\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{A}}})\det(\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}}) = 0$ since we assume that $\det(\boldsymbol{R}_{\boldsymbol{\Lambda}_{\boldsymbol{B}}}) = 0$.

Clearly, $Q(z)$ has no trivial roots, i.e., any known or unknown DOA are solution of $Q(z) = 0$. Inversely, we only have $I(z) = 0$ for the unknown DOA. So, according to (77), zeros of $\mathcal{C}(z)$ are only the DOA which annulate $I(z)$, i.e., the unknown DOA.

REFERENCES

[1] R. T. Behrens and L. L. Scharf, "Signal processing applications of oblique projection operators," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1413–1424, Jun. 1994.

[2] Y. Bistritz, "Zero location with respect to the unit circle of discrete-time linear system polynomials," *Proc. IEEE*, vol. 72, no. 9, pp. 1131–1142, Sep. 1984.

[3] G. Bouleux and R. Boyer, "Zero-forcing based sequential MUSIC algorithm," in *Proc. 32nd IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'07)*, 2007, vol. 3, pp. 1017–1020.

[4] R. Boyer and G. Bouleux, "Cramér-Rao lower bound for prior-subspace estimation," in *Proc. IEEE Sens. Array Multi-channel Process. (SAM 2006)*, 2006, pp. 394–398.

[5] H. Chen, S. Van Huffel, D. van Ormondt, and R. de Beer, "Parameter estimation with prior knowledge of known signal poles for the quantification of NMR spectroscopy data in the time domain," *J. Magnet. Reson.*, ser. A, vol. 119, no. 2, pp. 225–234, Apr. 1996.

[6] R. D. DeGroat, E. M. Dowling, and D. A. Linebarger, "The constrained MUSIC problem," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1445–1449, Mar. 1993.

[7] Y. Ephraim, N. Merhav, and H. L. Van Trees, "Min-norm interpretations and consistency of MUSIC, MODE and ML," *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2937–2942, Dec. 1995.

[8] P. C. Hansen and S. H. Jensen, "Prewhitening for rank-deficient noise in subspace methods for noise reduction," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pt. 1, pp. 3718–3726, Oct. 2005.

[9] S. Kayalar and H. L. Weinert, "Oblique projections: Formulas, algorithms and error bounds," *Math. Contr. Signals Syst.*, vol. 2, no. 1, pp. 33–45, 1989.

[10] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[11] D. A. Linebarger, R. D. DeGroat, E. M. Dowling, P. Stoica, and G. L. Fudge, "Incorporating *a priori* information into MUSIC-algorithms and analysis," *Signal Process.*, vol. 46, no. 1, pp. 84–105, Sep. 1995.

[12] M. L. McCloud and L. L. Scharf, "A new subspace identification algorithm for high-resolution DOA estimation," *IEEE Trans. Antennas Propag.*, vol. 50, no. 10, pp. 1382–1390, Oct. 2002.

[13] J. C. Mosher and R. M. Leahy, "Source localization using recursively applied and projected (RAP) MUSIC," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 332–340, Feb. 1999.

[14] S. K. Oh and C. K. Un, "A sequential estimation approach for performance improvement of eigenstructure-based methods in array processing," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 457–463, Jan. 1993.

[15] B. D. Rao and K. V. S. Hari, "Performance analysis of root-music," *IEEE Trans. Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.

[16] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1981.

[17] A. A. Shah and D. W. Tufts, "Determination of the dimension of a signal subspace from short data records," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2531–2535, Sep. 1994.

[18] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Englewood Cliffs, NJ: Prentice-Hall, 2005.

[19] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Trans. Signal Process.*, vol. 39, no. 2, pp. 446–453, Feb. 1991.

[20] P. Stoica and A. Nehorai, "Comparative performance study of element-space and beamspace MUSIC estimators," *Circuits, Syst. Signal Process.*, vol. 10, no. 3, pp. 285–292, 1991.

[21] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[22] P. Stoica, Y. Selen, N. Sandgren, and S. Van Huffel, "Using prior knowledge in SVD-based parameter estimation for magnetic resonance spectroscopy—the ATP example," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1568–1578, Sep. 2004.

[23] D. Thubert and L. Kopp, "Measurement accuracy and resolving power of high resolution passive methods," *Proc. EURASIP*, vol. 86, pp. 1037–1040, 1986.
[24] P. Vandaele and M. Moonen, "Two deterministic blind channel estimation algorithms based on oblique projections," *Signal Process.*, vol. 80, no. 3, pp. 481–495, Mar. 2000.
[25] A.-J. Van Der Veen, E. F. Deprettere, and A. L. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," *Proc. IEEE*, vol. 81, no. 9, pp. 1277–1308, Sep. 1993.
[26] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1110–1121, May 1991.
[27] P. Wirawan, K. Abed-Meraim, H. Maitre, and P. Duhamel, "Blind multichannel image restoration using oblique projections," in *Proc. IEEE Sens. Array Multichannel Process. (SAM2002)*, 2002, pp. 125–129.
[28] X. Yu and L. Tong, "Joint channel and symbol estimation by oblique projections," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3074–3083, Dec. 2001.

Rémy Boyer received the M.Sc. and Ph.D. degrees in the field of signal processing and communications from École Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1999 and 2002, respectively.

From 2002 to 2003, he was a Visiting Researcher with the University of Sherbrooke, QC, Canada. He is currently an Associate Professor with the Laboratoire des Signaux et Systèmes (L2S), Gif-sur-Yvette , with the CNRS-UPS-SUPELEC (National Center of Scientific Research and University of Paris-Sud). His research activities include multiway-arrays processing, sensor array, sinusoidal modeling, and watermarking.

Guillaume Bouleux received the engineering degree in electronic and signal processing from the Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA), Cergy, France, in 2004 and the Master degree from Universitée de Cergy-Pontoise, France, in 2005.

He is currently working toward the Ph.D. degree in signal processing in collaboration with both Universitée Jean Monnet (to LASPI laboratory), Saint Etienne, France and Universitée de Paris-Sud, Orsay, France (to LSS-Supelec laboratory). His research interest includes array signal processing, statistical analysis, linear, and multilinear algebra.